

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
14. Dezember 2000 (14.12.2000)

PCT

(10) Internationale Veröffentlichungsnummer  
**WO 00/75368 A2**

- (51) Internationale Patentklassifikation<sup>7</sup>: C12Q 1/68, C12N 15/10, A61K 48/00 (74) Anwälte: BETTENHAUSEN, Berthold usw.; Müllerstrasse 1, D-80469 München (DE).
- (21) Internationales Aktenzeichen: PCT/DE00/01863 (81) Bestimmungsstaaten (*national*): DE, JP, US.
- (22) Internationales Anmeldedatum: 7. Juni 2000 (07.06.2000) (84) Bestimmungsstaaten (*regional*): europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität: 199 25 862.7 7. Juni 1999 (07.06.1999) DE
- (71) Anmelder (*für alle Bestimmungsstaaten mit Ausnahme von US*): DIAVIR GMBH [DE/DE]; Steinbergstr. 30, D-85250 Altomünster (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (*nur für US*): SCHATZ, Octavian [DE/DE]; Steinbergstr. 30, D-85250 Altomünster (DE).
- Veröffentlicht: — Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.
- Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: METHOD FOR THE SYNTHESIS OF DNA FRAGMENTS

(54) Bezeichnung: VERFAHREN ZUR SYNTHESE VON DNA-FRAGMENTEN

(57) Abstract: The invention relates to a method that can be carried out in parallel and automated for the production of any nucleic acid, comprising the following steps: a) coupling an oligonucleotide to a solid matrix; b) adding an additional oligonucleotide; c) performing ligation of the oligonucleotide from steps a) and b) in an orientation; d) removing excess reactants and enzymes from the reaction preparation; e) effecting cleavage of the ligation product from step c) with a restriction system that cleaves outside the recognition sequence, whereby cleavage is effected in the shortened or lengthened oligonucleotide from step a) or in the oligonucleotide from step b); f) separating the reaction mixture from the lengthened or shortened oligonucleotide from step a); g) repeating at least one steps b) to f); h) performing successive sequence-independent linkage of the fragments obtained after executing steps a) to g) until the desired product is obtained.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft ein parallel ausführbares und automatisierbares Verfahren zur Herstellung einer beliebigen Nukleinsäure, umfassend die Schritte: a) Kopplung eines Oligonukleotids an eine feste Matrix; b) Zugabe eines weiteren Oligonukleotids; c) Ligation der Oligonukleotide aus Schritt a) und b) in einer Orientierung; d) Entfernung überschüssiger Reaktanden und Enzyme aus dem Reaktionsansatz; e) Spaltung des Ligationsprodukts aus Schritt c) mit einem ausserhalb der Erkennungssequenz spaltenden Restriktionsenzym, wobei die Spaltung entweder im Oligonukleotid aus Schritt a) oder im Oligonukleotid aus Schritt b) stattfindet; f) Abtrennen des Reaktionsgemisches vom in Schritt e) erhaltenen verlängerten oder verkürzten Oligonukleotid aus Schritt a); g) mindestens einmaliges Wiederholen der Schritte b) bis f); h) sukzessive sequenzunabhängige Verknüpfung der nach Durchführung der Schritte a) bis g) erhaltenen Fragmente bis zum Erhalt des gewünschten Produkts.

WO 00/75368 A2

5

## Verfahren zur Synthese von DNA-Fragmenten

Nach dem derzeitigen Stand der Technik müssen für eine Synthese einer etwa 2,5 kb großen  
10 Nukleinsäuresequenz zunächst etwa 50 verschiedene, teilweise überlappende ca. 80mer Oligo-  
nukleotide synthetisiert und aufgereinigt werden. Diese werden dann paarweise oder in Subsets  
hybridisiert und mittels einer Klenow-Polymerase-Reaktion aufgefüllt oder mit den außen  
liegenden Oligonukleotiden als Primer in einer Polymerasekettenreaktion (PCR) hergestellt und  
(meist über einzubauende Restriktionsstellen) unidirektional miteinander verknüpft. Dieses  
15 Verfahren ist als "Gap filling" Methode bekannt. Alternativ lassen sich Genfragmente durch  
enzymatische oder chemische Ligation synthetisieren; diese Fragmente können dann nach  
Aufreinigung und/oder Klonierung zu größeren Genabschnitten zusammengesetzt werden (sog.  
Cassettenmethode). Beide Prozeduren erfordern im Idealfall mindestens eine Woche, im Regel-  
fall jedoch eher 6-12 Wochen, mitunter sogar 6 Monate. Sequenzielle, an Festphasen gebundene  
20 Verfahren haben wegen der Vielzahl der notwendigen Reaktionsschritte nur geringe Ausbeuten  
und sind daher auch sehr fehleranfällig.

Eines der Hauptprobleme besteht darin, daß längere Oligonukleotide aus Gründen der  
Kopplungseffizienz, die selbst bei gut verlaufenden Synthesen nur 99% pro Schritt erreicht,  
25 immer einen unvermeidbaren Anteil an Abbruchprodukten aufweisen. Darüber hinaus kommt es  
auch zu Deletionen, die aus nicht 100%igem Capping resultieren. Selbst bei sehr guten  
Synthesen liegt dieser Anteil bei ca. 0,25% pro Kopplungsschritt. Auch die Abtrennung der  
Tritylschutzgruppen nach Beendigung der Synthese verläuft nicht vollständig. Die so  
entstandenen unvollständigen Oligonukleotidprodukte können selbst mit großem Aufwand von  
30 längeren Oligonukleotiden nicht vollständig abgetrennt werden.

Bei einer durchschnittlichen Kopplungseffizienz von 98% erhält man beispielsweise bei einem  
80mer eine Ausbeute des gewünschten Produkts mit voller Länge von lediglich 19.86%. Mit den  
heutzutage verfügbaren Aufreinigungsverfahren kann das gewünschte Endprodukt in einer

Reinheit von bestenfalls 95% dargestellt werden. Auch wenn dann nur noch ein kleiner Teil der endgereinigten Oligonukleotide fehlerhaft ist, steigt dennoch die Wahrscheinlichkeit einer fehlerhaften Endsequenz mit der Zahl der eingesetzten Oligonukleotide dramatisch an. Eine Sequenz, die sich aus 50 der beschriebenen Oligonukleotide zusammensetzt, ist demnach nur in 7,7% aller  
5 Fälle korrekt und muss deshalb in aller Regel nachgearbeitet werden. Ein relativ seltener Einbau falscher Basen aufgrund von Fehlkopplungen während der Synthese ist dabei nicht berücksichtigt.

Auf Grund der Vielfalt möglicher Sequenzen selbst relativ kurzer Oligonukleotide (bereits von einem 30mer existieren über  $10^{18}$  mögliche Sequenzvarianten) ist es zudem praktisch unmöglich, Oligonukleotide für verschiedene Genkonstrukte wieder zu verwenden. Daher ist es technisch nicht machbar, sämtliche zur Generierung beliebiger Sequenzen benötigte Oligonukleotide vorzuhalten. Für jedes neue Genkonstrukt müssen jeweils neue Oligonukleotide synthetisiert und aufgereinigt werden. Nur ein Bruchteil des synthetisierten Materials wird jedoch tatsächlich für  
15 die Gensynthese eingesetzt, der Rest kann aus oben beschriebenen Gründen nicht verwertet werden. Die nicht gelöste Einbindung der Oligonukleotidsynthese und deren Aufreinigung in den Gensyntheseprozess ist eines der Haupthindernisse, weswegen eine vollständige Automatisierung dieses Prozesses derzeit technisch nur äußerst schwierig und praktisch wahrscheinlich überhaupt nicht zu realisieren.

20

Die der vorliegenden Erfindung zugrundeliegende Aufgabe besteht somit in der Bereitstellung eines Verfahrens zur effizienten Synthese doppelsträngiger DNA-Fragmente beliebiger Sequenz und Länge. Eine weitere Aufgabe besteht darin, eine Methode bereit zu stellen, welche es erlaubt, beliebige DNA-Moleküle aus einer begrenzten Bibliothek von Grundbausteinen  
25 zusammen zu setzen. Eine weitere Aufgabe besteht darin, ein Verfahren aufzuzeigen, welches die **parallele** Synthese und die **sequenzunabhängige** Verknüpfung beliebiger Genfragmente gestattet. Die Erfüllung dieser beiden Voraussetzungen ist notwendig für die Realisierung einer kompletten Automatisierung des Gensyntheseverfahrens. Eine weitere Aufgabe besteht in der Bereitstellung eines Kits zur automatisierten Herstellung doppelsträngiger DNA-Fragmente.

30

Die Aufgabe wird gelöst durch die Bereitstellung eines Verfahrens zur Herstellung eines Nukleinsäuremoleküls, umfassend die Schritte:

- a) Kopplung eines Oligonukleotids mit einem Ende an eine feste Matrix, wobei die Kopplung über eine Modifikation erfolgt, und das Oligonukleotid eine Erkennungssequenz für ein TypII Restriktionsenzym enthält, welches außerhalb seiner Erkennungssequenz schneidet,
- 5 b) Zugabe eines weiteren, mindestens teilweise doppelsträngigen Oligonukleotids mit einer anderen Erkennungssequenz für ein TypII Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, als in Schritt a), wobei dieses Oligonukleotid nicht an die Matrix binden kann,
- c) Ligation der Oligonukleotide aus Schritt a) und b) in der durch die Blockierung der  
10 nicht zu ligierenden Enden vorgegebenen Orientierung,
- d) Entfernen nicht verbrauchter Reaktanden sowie Enzyme,
- e) Spaltung des Ligationsprodukts aus Schritt c) mit einem TypII Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung im Oligonukleotid aus Schritt a) stattfindet,
- 15 f) Abtrennen des erhaltenen Nukleinsäuremoleküls vom Reaktionsgemisch.

Die Aufgabe wird ferner gelöst durch die Bereitstellung eines Verfahrens zur Herstellung eines Nukleinsäuremoleküls, umfassend die Schritte:

- a) bis d) wie oben,
- 20 e) Spaltung des Ligationsprodukts aus Schritt c) mit einem TypII Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung in der Nukleinsäuresequenz des Oligonukleotids aus Schritt b) stattfindet,
- f) Abtrennen des Reaktionsgemisches vom in Schritt e) erhaltenen verlängerten Oligonukleotid aus Schritt a),
- 25 g) mindestens einmaliges Wiederholen der Schritte b) bis f).

Die Aufgabe wird ferner gelöst durch die Bereitstellung eines Verfahrens zur Herstellung eines Nukleinsäuremoleküls, umfassend die Schritte:

- a) bis g) wie oben,
- 30 h) Spaltung des erhaltenen Nukleinsäuremoleküls mit einem TypII Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung im Oligonukleotid aus Schritt a) stattfindet, und gegebenenfalls

- i) Spaltung des erhaltenen Nukleinsäuremoleküls mit einem TypII-S Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung im Oligonukleotid aus Schritt b) stattfindet.

5 Bevorzugt ist ein Verfahren, wobei nach Schritt c) als Schritt c)' eine Exonuklease- und/oder Phosphatase-Reaktion durchgeführt wird. Ferner bevorzugt ist ein Verfahren, wobei das Reaktionsgemisch des Schrittes c)' nach der Reaktion entfernt wird. Ferner bevorzugt ist ein Verfahren, wobei in der letzten Wiederholung der Schritte b) bis f) der Schritt e) nicht durchgeführt wird. Ferner bevorzugt ist ein Verfahren, wobei die erhaltene Nukleinsäure durch

10 Restriktionsspaltung vom Oligonukleotid aus Schritt a) abgetrennt wird. Ferner bevorzugt ist ein Verfahren, wobei die Kopplung des Oligonukleotids aus Schritt a) an die feste Matrix über eine Modifikation erfolgt. Besonders bevorzugt ist ein Verfahren, wobei die Modifikation ein Biotinrest, ein Digoxigeninrest, ein Fluoresceinisothiocyanat (FITC), eine Aminoverbindung oder ein Succinylester ist. Ferner bevorzugt ist ein Verfahren, wobei das Oligonukleotid aus

15 Schritt a) und/oder b) über einen Loop verfügt. Besonders bevorzugt ist ein Verfahren, wobei das Oligonukleotid aus Schritt a) mit dem Loop an die feste Matrix gekoppelt ist. Besonders bevorzugt ist ein Verfahren, wobei die feste Matrix ein Kügelchen (bead), vorzugsweise aus Glas oder Polystyrol, ein Objektträger, ein DNA Chip, die Vertiefung einer Mikrotiterplatte (well) oder ein Reaktionsröhrchen ist. Insbesondere bevorzugt ist ein Verfahren, wobei die feste Matrix

20 einen Streptavidinrest, einen anti-Digoxigenin-Antikörper oder einen anti-FITC-Antikörper umfasst. Ferner bevorzugt ist ein Verfahren, wobei die Oligonukleotide aus Schritt a) und b) an ihren zu ligierenden Enden über zueinander komplementäre Einzelstrangüberhänge verfügen. Besonders bevorzugt ist ein Verfahren, wobei die Einzelstrangüberhänge 1, 2, 3, 4 oder 5 Nukleotide lang sind. Besonders bevorzugt ist ein Verfahren, wobei die synthetisierte

25 Nukleinsäure in einem abschließenden Schritt mit einer replikationsfähigen DNA (einem Plasmidvektor, einer Phagen- oder Virus-DNA, einem artifiziellen Chromosom, einem PCR Produkt oder einer weiteren künstlich hergestellten DNA verknüpft wird. Besonders bevorzugt ist ein Verfahren zur Herstellung Codon-optimierter offener Leseraster, zur gezielten Mutagenese von Promotoren, Enhancern oder DNAs, welche für Proteine codieren. Insbesondere bevorzugt

30 ist die Verwendung der erfindungsgemäßen Nukleinsäure als Codon-optimierte DNA-Vakzine, zur Mutationsanalyse von Proteindomänen, als Matrize für Designerproteine, als Expressionskonstrukt für *in vitro* Proteinsynthese, zur Herstellung von Ribozymen oder



Aptameren, als Sonde zum Nachweis pathogener Mikroorganismen, als Sonde zum Nachweis der Expression von Genen, zum Nachweis allelspezifischer Mutationen, zum Nachweis von Protein/Protein-Bindung, Protein/Peptid-Bindung und/oder der Bindung niedermolekularer Stoffe an Proteine.

5

Die Aufgabe wird ferner gelöst durch die Bereitstellung eines Kits zur Herstellung einer Nukleinsäure nach dem erfindungsgemäßen Verfahren, umfassend

- 10 a) eine Bibliothek von 1 bis 1.048.576 verschiedenen Oligonukleotiden enthält, wobei die Oligonukleotide über eine Modifikation an einem Ende an eine feste Matrix koppelbar sind und das Oligonukleotid eine Erkennungssequenz oder einen Teil der Erkennungssequenz für ein TypII Restriktionsenzym enthält, welches außerhalb seiner Erkennungssequenz schneidet,
- 15 b) eine weitere Bibliothek von 4 bis 1.048.576 verschiedenen Oligonukleotiden enthält, wobei jedes der Oligonukleotide eine Erkennungssequenz für ein TypII Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, enthält, dass von dem TypII Restriktionsenzym aus a) verschieden ist, und gegebenenfalls den anderen Teil der Erkennungssequenz des Restriktionsenzym aus Schritt a) enthält,
- 20 c) eine feste Matrix,
- d) Reservoirs der für die Herstellung des Nukleinsäuremoleküls benötigten Enzyme und/oder anderer Reagenzien.

Bevorzugt ist ein Kit, wobei die Enzyme eine Ligase oder Topoisomerase und/oder ein oder mehrere Restriktionsenzym(e) und/oder eine Exonuklease und/oder eine Phosphatase umfassen.

- 25 Besonders bevorzugt ist ein Automat, welcher nach Eingabe der gewünschten Basensequenz sämtliche Reaktionsschritte fest legen und selbsttätig abarbeiten kann.

Die Erfindung wird weiter durch die folgenden Figuren erläutert.

- 30 Figur 1 zeigt eine schematische Darstellung des erfindungsgemäßen Verfahrens. Bio bedeutet eine Modifikation (z.B. Biotin), mit der das Anchor-Oligonukleotid an eine feste Matrix (z.B. Streptavidin) gekoppelt ist. T, G, C, A und N bezeichnen die Nukleinsäurebasen, wobei T

Thymin, G Guanin, C Cytosin, A Adenin und N eine beliebige der vier Nukleinsäurebasen bedeutet.

Figur 2 zeigt schematisch den Aufbau eines EasyPro™ Transkriptions/Translationssystems von PCR-Fragmenten. Bio bedeutet eine Modifikation, mit der das Anchor-Oligonukleotid an eine feste Matrix gekoppelt ist. 5'-UTR bedeutet 5'-untranslatierter Bereich. ATG bedeutet Startcodon. 6 x His bedeutet eine sechsmalige Aneinanderreihung von Histidincodons. Single T overhang bedeutet einen Überhang von einem Thyminrest.

Figur 3 zeigt eine schematische Darstellung eines Minireaktors für die Proteinsynthese.

Figur 4 zeigt eine schematische Darstellung der Produktion einer Peptidlibrary mit dem QuickPep™-Verfahren. Bio bedeutet eine Modifikation, mit der das Anchor-Oligonukleotid an eine feste Matrix gekoppelt ist. T7 bedeutet T7-Promotor. rbs bedeutet interne Ribosomenbindungsstelle. ATG bedeutet Startcodon. EK bedeutet Enterokinase-Schnittstelle. Peptid ORF bedeutet den offenen Leserahmen des Peptids. STOP bedeutet das Stopcodon. Poly A bezeichnet den poly-A-Schwanz.

Figur 5 zeigt eine schematische Darstellung der Selektion von Ribozymen mit dem RiboSelect™-Verfahren.

Figur 6 zeigt eine schematische Darstellung des Nachweises von Pathogenen nach Anreicherung durch PCR (PathoCheck™). Bio bedeutet eine Modifikation, mit der das Anchor-Oligonukleotid an eine feste Matrix gekoppelt ist.

Figur 7 zeigt eine schematische Darstellung der Identifizierung bekannter Allele durch Ligation markierter Splinker (LIMA™). Bio bedeutet eine Modifikation, mit der das Anchor-Oligonukleotid an eine feste Matrix gekoppelt ist. x bedeutet die Stelle, an der die zu bestimmende Modifikation vorliegt.

Figur 8 zeigt eine schematische Darstellung der Parallelanalyse von mRNA-Arrays (PAMINA™).

Figur 9 zeigt die schematische Darstellung eines Anchor-Oligonukleotids. Bio bedeutet eine Modifikation, mit der das Anchor-Oligonukleotid an eine feste Matrix gekoppelt ist. T, G, C, A bezeichnen die Nukleinsäurebasen, wobei T Thymidin, G Guanin, C Cytosin, A Adenin  
5 bedeutet. Esp3I bezeichnet ein Restriktionsenzym.

Figur 10 zeigt die schematische Darstellung eines Anchor-Oligonukleotids. Bio bedeutet eine Modifikation, mit der das Anchor-Oligonukleotid an eine feste Matrix gekoppelt ist. T, G, C, A bezeichnen die Nukleinsäurebasen, wobei T Thymidin, G Guanin, C Cytosin, A Adenin  
10 bedeutet. BpiI bezeichnet ein Restriktionsenzym.

Figur 11 zeigt die schematische Darstellung eines Bipartite-Anchor-Oligonukleotids. Bio bedeutet eine Modifikation, mit der das Anchor-Oligonukleotid an eine feste Matrix gekoppelt ist. T, G, C, A bezeichnen die Nukleinsäurebasen, wobei T Thymidin, G Guanin, C Cytosin, A  
15 Adenin bedeutet.

Figur 12 zeigt die schematische Darstellung eines Splinker-Oligonukleotids. T, G, C, A bezeichnen die Nukleinsäurebasen, wobei T Thymidin, G Guanin, C Cytosin, A Adenin bedeutet. BsaI und Eco31I bezeichnen Restriktionsenzyme.  
20

Figur 13 zeigt die schematische Darstellung eines Bipartite-Splinker-Oligonukleotids. T, G, C, A bezeichnen die Nukleinsäurebasen, wobei T Thymidin, G Guanin, C Cytosin, A Adenin bedeutet. BsaI und Eco31I bezeichnen Restriktionsenzyme.

25 Figur 14 zeigt eine schematische Darstellung des Synthesewegs längerer Nukleinsäuren mit dem erfindungsgemäßen Verfahren. Die Balken symbolisieren doppelsträngige DNA-Fragmente, welche parallel durch hintereinander geschaltete Ligations/Restriktionszyklen synthetisiert wurden. Im Endprodukt benachbarte Teilabschnitte werden jeweils durch die Ligation eines aufligierten Splinkers mit einem aufligierten Anchor verbunden. Die so gewonnenen größeren  
30 Fragmente werden dann im nächsten Schritt wieder entweder mit der Anchor-spezifischen oder der Splinker-spezifischen Restriktionsendonuklease geschnitten und über komplementäre Überhänge miteinander verknüpft usw., so dass sich die Länge der Fragmente mit jedem Schritt



verdoppelt. Die Verknüpfung ist vollkommen sequenzunabhängig, da die Erkennungssequenzen der verwendeten Restriktionsendonukleasen in den jeweils abgeschnittenen Teilen der aufligierten Fragmente liegen und daher nicht in die wachsende Nukleinsäure eingebaut werden. Die Zahlen über den Balken bedeuten die Größe der Fragmente in Basenpaaren. Ausgehend von 20 Basenpaar großen DNA-Fragmenten ergibt sich somit nach vier Transpositionen eine max. Länge von 320 Basenpaaren, nach fünf Transpositionen eine Länge von 640 Basenpaaren, nach sechs Transpositionen eine Länge von 1280 Basenpaaren, nach sieben Transpositionen eine Länge von 2560 Basenpaaren, usw.

## 10 Definitionen

Der hier verwendete Begriff "parallel" oder "parallele Synthese" bedeutet, daß verschiedene erfindungsgemäße Nukleinsäuremoleküle gleichzeitig in getrennten Reaktionsansätzen synthetisiert werden können, um dann mit dem erfindungsgemäßen Verfahren z.B. als Anchor oder Splinker zu einem verlängerten Nukleinsäuremoleküle ligiert werden können.

Der hier verwendete Begriff "Sloning" (Sequentielle Ligation von Oligonukleotiden auf Sequenz-unabhängige Weise) bezieht sich auf ein Verfahren zur aufeinanderfolgenden Ligation von Oligonukleotiden mit beliebiger Sequenz.

20

Der hier verwendete Begriff "Anchor" oder "Anchor-Oligonukleotid" bezieht sich auf ein Oligonukleotid, welches über eine Modifikation an eine feste Matrix gekoppelt werden kann. Im Sinne der vorliegenden Erfindung enthält das Oligonukleotid in seinem doppelsträngigen Anteil ferner eine Restriktionsschnittstelle für ein TypIIIS Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet.

Der hier verwendete Begriff "Splinker" oder "Splinker-Oligonukleotid" bezieht sich auf ein Oligonukleotid, welches über keine bzw. eine andersgeartete Modifikation verfügt, so daß es nicht selbst an die Matrix bindet, an welche die Anchor-Oligonukleotide gekoppelt sind.

30

Der hier verwendete Begriff "Dumbbell" (auf dt.: Glockenklöppel) bezieht sich auf eine DNA-Struktur, die durch einen Doppelstrang charakterisiert ist, der von zwei Loops flankiert wird.

Ein Aspekt der vorliegenden Erfindung betrifft ein Verfahren zur Herstellung eines Nukleinsäuremoleküls, umfassend die Schritte:

- 5 a) Kopplung eines Oligonukleotids mit einem Ende an eine feste Matrix, wobei die Kopplung über eine Modifikation erfolgt, und das Oligonukleotid eine Erkennungssequenz für ein TypIIS Restriktionsenzym enthält, welches außerhalb seiner Erkennungssequenz schneidet,
- 10 b) Zugabe eines weiteren, mindestens teilweise doppelsträngigen Oligonukleotids mit einer anderen Erkennungssequenz für ein TypIIS Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, als in Schritt a), wobei dieses Oligonukleotid nicht an die Matrix binden kann,
- d) Ligation der Oligonukleotide aus Schritt a) und b) in der durch die Blockierung der nicht zu ligierenden Enden vorgegebenen Orientierung,
- h) Entfernen nicht verbrauchter Reaktanden sowie Enzyme,
- 15 i) Spaltung des Ligationsprodukts aus Schritt c) mit einem TypIIS Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung im Oligonukleotid aus Schritt a) stattfindet,
- j) Abtrennen des erhaltenen Nukleinsäuremoleküls vom Reaktionsgemisch.

Ein weiteren Aspekt der Erfindung betrifft ein Verfahren zur Herstellung eines Nukleinsäuremoleküls, umfassend die Schritte:

- a) bis d) wie oben,
- e) Spaltung des Ligationsprodukts aus Schritt c) mit einem TypIIS Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung in der Nukleinsäuresequenz des Oligonukleotids aus Schritt b) stattfindet,
- 25 f) Abtrennen des Reaktionsgemisches vom in Schritt e) erhaltenen verlängerten Oligonukleotid aus Schritt a),
- k) mindestens einmaliges Wiederholen der Schritte b) bis f).

Ein weiteren Aspekt der Erfindung betrifft ein Verfahren zur Herstellung eines Nukleinsäuremoleküls, umfassend die Schritte:

- a) bis g) wie oben,
- h) Spaltung des erhaltenen Nukleinsäuremoleküls mit einem TypIIS Restriktionsenzym,

welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung im Oligonukleotid aus Schritt a) stattfindet, und gegebenenfalls

- 5 i) Spaltung des erhaltenen Nukleinsäuremoleküls mit einem TypIIIS Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung im Oligonukleotid aus Schritt b) stattfindet.

Eines der beiden in jedem Reaktionsschritt zu verknüpfenden Oligonukleotide (das sog. "Anchor"-Oligonukleotid) ist über eine Modifikation, z.B. eine niedermolekulare chemische Verbindung wie Biotin oder Digoxigenin, an eine feste Matrix koppelbar. In einer bevorzugten Ausführungsform handelt es sich dabei um magnetische Streptavidin-beschichtete oder anti-Digoxigenin-beschichtete Beads. Das andere Oligonukleotid (das sog. "Splinker"-Oligonukleotid) besitzt auch ein blockiertes Ende, aber keine derartige oder aber eine andersartige Modifikation. Der Punkt auf den es ankommt ist der, dass die Anchor-Oligonukleotide durch die Bindung an eine geeignete Matrix von den Splinker-Oligonukleotiden getrennt werden können. Es können daher beliebige Verbindungen z.B. Biotin, Digoxigenin, 10 Fluoresceinisothiocyanat (FITC), Aminoverbindungen, Succinylester und andere dem Fachmann geläufige Verbindungen, verwendet werden, sofern sie geeignet sind, direkt oder indirekt (z.B. über einen Antikörper) eine Bindung an eine feste Phase zu vermitteln.

20 Anchor-Oligonukleotide können entweder aus einem einzigen zum Teil selbstkomplementären Oligonukleotid bestehen, welches über eine Modifikation vorzugsweise in der Loopsequenz an eine feste Phase koppelbar ist, oder aus zwei einzelsträngigen Oligonukleotiden, welche einen Doppelstrang bilden, vorzugsweise mit einem Einzelstrangüberhang. Da nur einer der beiden Stränge an eine Matrix gekoppelt werden muss, kann der andere wenn nötig durch Alkali oder 25 Hitze denaturiert und abgetrennt werden (etwa um als Template für eine PCR-Reaktion zu dienen). Um sicherzugehen, daß auch bei solchen zweigeteilten Anchor-Oligonukleotiden nur ein Ende ligierbar ist, werden die nicht für die Ligation benötigten Enden entsprechend blockiert. Nukleinsäuresequenzen beispielhafter Anchor-Oligonukleotide sind

30 Anchor A3I

5'-GCTTCGAGACGCGTTTTTCGCGTCTCG-3' (SEQ ID NR:1; FIG. 9)

## Anchor A2+

5'-AGAATGGTCTTCGAGCTTTTGCTCGAAGACCA-3' (SEQ ID NR:2; FIG. 10)

## Bipartite Anchor

5 5'-CGCGGATCCGCGGCGT-3' (SEQ ID NR:3, FIG. 11)

5'-CGAGACGCCGCGGATCCGCG-3' (SEQ ID NR:4, FIG. 11)

Splinker-Oligonukleotide können entweder aus einem einzigen zum Teil selbstkomplementären Oligonukleotid bestehen oder aus zwei einzelsträngigen Oligonukleotiden, welche einen  
 10 Doppelstrang bilden, vorzugsweise mit einem Einzelstrangüberhang, d.h. man hat ein mindestens teilweise komplementäres Paar von Oligonukleotiden, wobei die jeweils nicht zu ligierenden Enden der beiden Einzelstränge blockiert sein müssen. Die bevorzugte Einzelstrangüberhangsequenz muss zu dem jeweils zu ligierenden Anchor-Oligonukleotid komplementär sein. Nukleinsäuresequenzen beispielhafter Splinker-Oligonukleotide sind

15

## Splinker S1H

5'-AAGCTTCTGGAGACCGCTTTTGCGGTCTCCAGAA-3' (SEQ ID NR:5, FIG. 12)

## Bipartite Splinker

20 5'-CTCGAAGCGGAGACCGCCAC-3' (SEQ ID NR:6, FIG. 13)

5'-GTGGCGGTCTCCGCTT-3' (SEQ ID NR:7, FIG. 13)

Sowohl Anchor- wie auch Splinker-Oligonukleotide können Überhänge einer definierten Länge enthalten, in einer bevorzugten Ausführungsform eins bis fünf Nukleotide. Diese Überhänge sind  
 25 bei den jeweils zu ligierenden Oligonukleotiden komplementär zueinander, am 5' Ende phosphoryliert und können nur in einer Orientierung miteinander ligiert werden. Dabei entsteht ein ligiertes Oligonukleotid mit z.B. einer sogenannten "Dumbbell"-Struktur. Um eine vollständige Ligation aller verfügbaren Anchor Oligonukleotide zu erreichen, können die anzuligierenden Splinker-Oligonukleotide in 2 bis 10fachem Überschuß zugesetzt werden. Die  
 30 überschüssigen, nicht-reagierten Splinker werden nach jedem Ligationsschritt mit Puffer gewaschen. Werden z.B. mit Streptavidin beschichtete magnetische Beads verwendet, können die Beads mit den über eine Streptavidin/Biotin-Bindung gebundenen Anchor-

Oligonukleotiden zusammen mit den anligierten Splinkern durch den Einsatz eines Magneten im Reaktionsansatz zurückgehalten werden. Alternativ können z.B. direkt mit Streptavidin beschichtete Wells, Glas(beads), Objektträger, DNA-Chips oder beliebige andere feste Phasen verwendet werden. Beads werden in der Regel bevorzugt, weil sie eine größere Oberfläche und  
5 daher eine höhere Bindungskapazität aufweisen.

Um weitere Ligationen durchführen zu können, muss eine Erkennungssequenz für eine Restriktionsendonuklease vorhanden sein, die die Nukleinsäuresequenz außerhalb dieser Erkennungssequenz in dem anligierten Splinker-Oligonukleotid schneidet. Beispiele für solche  
10 Enzyme sind BpiI, Esp3I, Eco31I, SapI etc. Für das erfindungsgemäße Verfahren nützliche Restriktionsenzyme und ihre Erkennungssequenzen und Schnittstellen sind in der Rebase-Datenbank unter <http://rebase.neb.com/rebase/rebase.html> zu finden. An der in den Splinker-Oligonukleotiden enthaltenen Restriktionsschnittstelle werden die Ligationsprodukte so geschnitten, daß ein Teil der Splinkersequenz am Anchor-Oligonukleotid verbleibt. Gleichzeitig  
15 wird dadurch ein Sequenzüberhang erzeugt, der für die Ligation eines weiteren Splinker-Oligonukleotids herangezogen werden kann. Der andere abgespaltene Teil und der nicht ligierten Reste des Splinker-Oligonukleotids, das Restriktionsenzym sowie der Restriktionspuffer werden aus dem Reaktionsansatz ausgewaschen, worauf ein weiterer Zyklus beginnt. Der Zyklus kann entweder nur einmal durchgeführt oder aber mehrere Male wiederholt werden, bevor die so  
20 verlängerten Oligonukleotide ihrerseits mit den parallel synthetisierten Nachbarfragmenten verknüpft werden. Da die beim Schneiden mit den verschiedenen Restriktionsendonukleasen entstehenden, jeweils zueinander komplementären Überhänge aus dem zu synthetisierenden Gen stammen, die Erkennungssequenzen hingegen in den wieder abgeschnittenen Teilen der Anchor- bzw. Splinker-Oligonukleotide lokalisiert sind, können die benachbarten Fragmente völlig unab-  
25 hängig von ihrer Sequenz verknüpft werden. Insbesondere lassen sich dadurch selbst große Gene in vielen parallel ablaufenden Teilreaktionen in nur wenigen Reaktionsschritten zusammen bauen. Im optimalen Fall lässt sich beispielsweise ein 2 kb großes Gen in lediglich 9 Schritten aus 256 Einzelreaktionen zusammen setzen. Ein gleich großes Gen würde bei linearer Synthese (rekursiv, aber nicht parallel) und bei Verwendung von 60mer Oligos mehr als 30 Schritte  
30 benötigen. Da sowohl enzymatische als auch chemische Ligationsverfahren Ausbeuten von meist nur 80-90% aufweisen, nimmt die Gesamtausbeute mit der Anzahl der benötigten Reaktionsschritte exponentiell ab, weswegen Verfahren mit wenigen Reaktionsschritten



- vorteilhaft sind. Um nicht-umgesetzte Anchor-Oligonukleotide von der weiteren Synthese auszuschließen, kann fakultativ nach der Ligation ein Exonuklease- und/oder Phosphatase-Schritt zwischengeschaltet werden, wodurch der Überhang oder zumindest die für die folgende Ligation erforderliche 5'-Phosphatgruppe entfernt wird. Der Anteil an nicht-umgesetzten Anchor-Oligonukleotiden ist bei einem eingesetzten Überschuß an Splinker-Oligonukleotiden nur gering. Eine Weiterreaktion sollte zudem nur dann möglich sein, wenn dieselbe Sequenz ein weiteres Mal anligiert wird, weswegen die Gefahr einer Kontamination mit nicht oder nur teilweise umgesetzten Anchor Oligonukleotiden als relativ gering anzusehen ist.
- Die so nach mehreren Ligations- und Restriktionszyklen aufligierte Nukleinsäuresequenz kann anschließend durch Schneiden mit einem Restriktionsenzym, das eine Nukleinsäuresequenz in dem ursprünglichen Anchor-Oligonukleotid spezifisch erkennt, von dem an der Matrix verbleibenden Anchor-Oligonukleotid abgetrennt werden. Die aufligierte Nukleinsäuresequenz hängt nun an dem zuletzt anligierten Splinker-Oligonukleotid. Das verlängerte Splinker-Oligonukleotid wird nach Inaktivierung des Restriktionsenzym aus dem ursprünglichen Reaktionsansatz in ein neues Reaktionsgefäß überführt und dort mit einem aufligierten Anchor-Oligonukleotid verknüpft, das mit einem Splinker-Oligonukleotid spezifischen Restriktionsenzym geschnitten wurde (1. Transposition). Dem Fachmann ist ersichtlich, daß es sich bei den aufligierten Nukleinsäuresequenzen um frei wählbare Sequenzen handeln kann, die sowohl unterschiedlich als auch identisch sein können. Das aus der 1. Transposition resultierende Ligationsprodukt wird wiederum mit einer Anchor-spezifischen Restriktionsendonuklease geschnitten und erneut mit einem analog erhaltenen aufligierten Anchor-Oligonukleotid ligiert (2. Transposition). Auf diese Weise verdoppelt sich die Länge der aufligierten Nukleinsäuresequenzen dann mit jedem weiteren Schritt. Die Verknüpfung der DNA-Fragmente erfolgt jeweils über komplementäre Überhänge, ist aber ansonsten vollkommen sequenzunabhängig. Die einzige Einschränkung dabei ist, daß die Anchor- und Splinker-spezifischen Restriktionsschnittstellen in der zu synthetisierenden Sequenz nicht vorkommen dürfen, weil sonst die DNA auch intern geschnitten werden würde. Jeweils vor einer Spaltung an einer Anchor-spezifischen Restriktionsschnittstelle und der darauf folgenden Transposition kann fakultativ ein Exonukleaseschritt eingeführt werden, um die Transposition nicht vollständig aufligierter Splinker-Oligonukleotide zu verhindern. Die für das Verfahren notwendigen sequenzspezifischen Spaltungen können statt durch TypIIIS Restriktionsendonukleasen im

Prinzip auch durch analog funktionierende Ribozyme erfolgen.

Ausgehend von einer 20 Basenpaar langen Sequenz (die man für Splinker mit einem 4 nt Überhang durch 5 sukzessive Ligationen der hierfür benötigten Ursprungs-Splinker aus der Bibliothek erhalten kann), lässt sich durch lediglich 7 weitere Ligationsschritte eine doppelsträngige DNA-Sequenz von 2560 Basenpaaren Länge synthetisieren. Bei Zykluszeiten von ca. 1 Stunde kann eine beliebige DNA-Sequenz dieser Länge binnen 12 Stunden synthetisiert werden. Durch eine Optimierung der Reaktionsbedingungen kann der Zeitaufwand auf etwa 6 Stunden halbiert werden.

10

Bei 4 Nukleotide langen Überhängen wird eine Bibliothek von 65536 verschiedenen Splinker-Oligonukleotiden benötigt, um alle möglichen Nukleinsäuresequenzen herstellen zu können. Diese Zahl ergibt sich aus folgender Berechnung: es gibt 256 mögliche 4 Nukleotide lange Überhänge ( $4^4 = 256$ ), ebenso viele Sequenzvarianten existieren für die vier direkt angrenzenden Nukleotide, die den Überhang beim nächsten Ligationsschritt bilden. Insgesamt ergibt sich daraus eine Gesamtzahl von  $4^4 \text{ mal } 4^4 = 4^8 = 65536$  Splinker-Oligonukleotide, mit denen sämtliche möglichen Sequenzvarianten dargestellt werden können. Bei 3 Nukleotide langen Überhängen reduziert sich die Komplexität der benötigten Splinkerbibliothek entsprechend auf  $4^3 \text{ mal } 4^3 = 4096$ , bei zwei Nukleotide langen Überhängen auf  $4^2 \text{ mal } 4^2 = 256$ , bei 5 Nukleotide langen Überhängen würde sie sich auf  $4^5 \text{ mal } 4^5 = 1048576$  erhöhen. Voraussetzung für dieses Baukastensystem ist das Vorhandensein einer kompletten Splinkerbibliothek (für 2 nt Überhänge 256 Oligonukleotide, für 3 nt Überhänge 4096 Oligonukleotide, für 4 nt Überhänge 65536 Oligonukleotide, für 5 nt Überhänge 1048576 Oligonukleotide) sowie einer Anchorbibliothek (4, 16, 64, 256 oder 1024 Oligonukleotide bei 1, 2, 3, 4 oder 5 nt Überhängen). Letztere ist allerdings nicht unbedingt nötig, da die verschiedenen Überhangsequenzen ebenso gut durch einen vorgeschalteten Ligationsschritt mit geeigneten Splinker-Oligonukleotiden erzeugt werden können.

25

Prinzipiell sind alle Einzelschritte des erfindungsgemäßen Verfahrens automatisierbar, so daß die Herstellung ganzer Gene so einfach ist wie die Synthese von Oligonukleotiden. Zudem eröffnet sich durch das erfindungsgemäße Verfahren ein Kostensenkungspotential in beträchtlicher Höhe. Erstens können alle benötigten Enzyme großtechnisch hergestellt werden. Zweitens können die

30

Investitionen für die Splinkerlibrary deutlich gesenkt werden, indem die einzelnen Splinker-Oligonukleotide bis auf die letzten 4 Nukleotide des 5'-Überhangs *en bloc* synthetisiert werden. Die Synthesereaktion wird dann in 4 gleiche Teile portioniert; die vier verschiedenen Nukleotide werden dann in separaten Reaktionen an der nächsten (im Endprodukt viertletzten) Position  
5 angehängt. Danach werden die vier Einzelreaktionen wiederum geviertelt, wonach das drittletzte Nukleotid angehängt wird usw. Statt 65536 Einzelsynthesen würde man dann nur 256 Synthesen in einem entsprechend größeren und deswegen günstigeren Maßstab benötigen. Ferner können die 256 möglichen 4 Nukleotide langen Überhänge durch eine "blunt end ligation" an 256 verschiedene Anchor-Oligonukleotide, anschließender Exonukleasebehandlung, Waschen und  
10 schließlich Restriktion mit der *Anchor*-spezifischen Restriktionsendonuklease erzeugt werden. Auf diese Weise könnten die 65536 benötigten Splinker-Oligonukleotide kostengünstig hergestellt werden. Ferner könnte auf diese Art und Weise eine aufwendige Aufreinigung aller 65536 Splinker-Oligonukleotide umgangen werden, da nichtreaktive Fehlsequenzen durch dieses Verfahren entfernt werden. Da eine extrem hohe Reinheit der eingesetzten Oligonukleotide  
15 essentiell für das Gelingen fehlerloser Synthesen ist, müssen diese sowieso entsprechend vorbehandelt werden. Daneben muss eine praktisch vollständige Abwesenheit von Exonukleasen während der Restriktions- und Ligationsschritte gewährleistet sein, damit die Überhangsequenzen intakt bleiben, die für die nachfolgenden Ligationen benötigt werden. Vor allem wenn Exonuklease-Zwischenschritte zur Entfernung nichtligierter Anchor-Oligonukleotide verwendet  
20 werden, müssen diese Exonukleasen gründlich gewaschen und/oder inaktiviert werden.

Die Anchor- und Splinker-Oligonukleotide können sowohl jeweils aus einem selbstkomplementären Einzelstrang als auch aus je zwei komplementären Plus- und Minus-Strängen zusammengesetzt werden. Die Nukleinsäuresequenzen müssen nicht vollständig komplementär  
25 sein; die selbstkomplementären Einzelstrang-Oligonukleotide können einen Loop aufweisen und die komplementären Plus- und Minus-Stränge können nur teilweise komplementär sein. Bei Anchor- und Splinker-Oligonukleotiden, die aus je zwei komplementären Plus- und Minus-Strängen zusammengesetzt sind, muss (i) die Schmelztemperatur des Doppelstranghybrids genügend hoch sein, um eine Denaturierung der zusammengesetzten Anchor- und Splinker-  
30 Oligonukleotide und einen möglicherweise daraus resultierenden unbeabsichtigten Transfer der nicht an eine Festphase gekoppelten Einzelstränge zu verhindern und müssen (ii) die jeweils nicht zu verlängernden Enden durch geeignete Modifikationen blockiert sein. Oligonukleotide

aus zwei komplementären Plus- und Minus-Strängen besitzen gegenüber Oligonukleotiden aus einem selbstkomplementären Einzelstrang bestimmte Vorteile. Selbstkomplementäre (Snap back) Oligonukleotide verursachen bei der Aufreinigung oft gewisse Schwierigkeiten, da sie in hoher Konzentration eine Tendenz zur Bildung von Netzwerken haben. Einzelsträngige Teil-  
 5 oligonukleotide sind auch kürzer und dadurch mit geringerem Aufwand in höherer Reinheit zu gewinnen. Für bestimmte erfindungsgemäße Ausführungsformen werden aus zwei Teil-oligonukleotiden zusammengesetzte ("bi-partite") Anchor-Oligonukleotide verwendet.

In besonders bevorzugten Ausführungsformen enthalten die Anchor bzw. Splinker die folgenden  
 10 Kombinationen an Erkennungssequenzen:

Anchor	Splinker
CGTCTCN^NNNN_ (Esp3I, BsmBI)	GGTCTCN^NNNN_ (BsaI, Eco31I,..)
GGTCTCN^NNNN_ (BsaI, Eco31I,..)	CGTCTCN^NNNN_ (Esp3I, BsmBI)
GAAGACNN^NNNN_ (BbsI, BpiI...)	ACCTGCNNNN^NNNN_ (BspMI, Acc36I)
ACCTGCNNNN^NNNN_ (BspMI, Acc36I)	GAAGACNN^NNNN_ (BbsI, BpiI...)
GCAGTG_NN^ (BtsI)	GCAATG_NN^ (BsrDI, Bse3DI, ..)
GCAATG_NN^ (BsrDI, Bse3DI, ..)	GCAGTG_NN^ (BtsI)
GTATCCNNNNN_N^ (BciVI, BfuI)	ACTGGGNNNN_N^ (BfiI, BmrI)
ACTGGGNNNN_N^ (BfiI, BmrI)	GTATCCNNNNN_N^ (BciVI, BfuI)

Ein weiterer Aspekt der vorliegenden Erfindung ist ein Kit zur Herstellung der Nukleinsäure nach dem erfindungsgemäßen Verfahren. Das Kit kann aus einer Bibliothek aller notwendigen  
 15 Anchor- und Splinker-Oligonukleotide, ferner einer festen Phase, an welche die Anchor-Oligonukleotide gekoppelt werden können, vorzugsweise magnetisierte Beads, geeigneten Reaktionsbehältnissen, Ligase, gegebenenfalls einer Topoisomerase und/oder einer 3'-5'-Exonuklease und/oder Phosphatase, mindestens zwei verschiedenen TypII-Restriktionsendonukleasen, die außerhalb ihrer Erkennungssequenz schneiden, sowie aller  
 20 benötigten Reaktionspuffer bestehen. Bevorzugt ist fernerhin eine Pipettierstation mit einem kühlbaren Probenvorratsbehälter mit einer entsprechenden Softwaresteuerung, die sämtliche Schritte des erfindungsgemäßen Verfahrens automatisch durchführt

Die vorliegende Erfindung erlaubt eine komplette Automatisierung des gesamten Gensynthese-  
 25 prozesses durch die Bereitstellung von einer Bibliothek immer wieder zu benützender,



mindestens teilweise doppelsträngiger hochreiner Oligonukleotide mit Erkennungssequenzen für bestimmte TypIIIS Restriktionsendonukleasen (sogenannte "Outside Cutter"). Die Automatisierung wird ferner erlaubt durch die Bereitstellung eines Verfahrens, welche die parallele Synthese von Genfragmenten und deren sequenzunabhängige Verknüpfung an jeder beliebigen Stelle erlaubt sowie aufgrund einer orientierungsgebundenen Verlängerung der Startmoleküle über deren Bindung an eine Festphase (die nicht zu ligierenden Enden sind durch geeignete Modifikationen bzw. Loopsequenzen blockiert) und eines definierten Sets rekursiver Prozeduren (Ligations-, Wasch- und Restriktionsschritte), die durch einen Roboter abgearbeitet werden können.

10

Nachstehend werden bestimmte Aspekte der vorliegenden Erfindung beispielhaft wiedergegeben, die auf der Komplettsynthese ganzer Gene durch das erfindungsgemäße Verfahren beruhen.

#### 1. Herstellung einer cDNA, wenn nur die Proteinsequenz bekannt ist

15 Es kommt häufig vor, daß nur die Aminosäuresequenz oder Teile der Aminosäuresequenz eines Proteins bekannt sind, jedoch nicht die cDNA oder genomische Sequenz. Wegen der Degeneration des genetischen Codes ist es in der Regel nicht möglich, das entsprechende Gen direkt über eine PCR einer geeigneten cDNA-Bank zu amplifizieren. Man sucht daher Regionen, in denen Aminosäuren wie Tryptophan, Methionin bzw. Asparagin, Aspartat, Glutamat, Glutamin, Tyrosin, Phenylalanin, Cystein oder Lysin gehäuft auftreten, da es für diese Aminosäuren nur ein bzw. zwei Codons gibt. Sofern es gelingt, mit niedrig degenerierten Primern ein PCR-Fragment der erwarteten Größe zu erhalten, wird dieses als Sonde benutzt, um das dazugehörige Gen aus einer cDNA-Bank zu klonieren. Diese Arbeit wird zwar heutzutage durch die Verfügbarkeit von Gene-Arrays und Klonkollektionen in vielen Fällen wesentlich erleichtert, doch zum einen stehen solche Hilfsmittel nur für eine begrenzte Anzahl von Organismen und Zelltypen zur Verfügung, zum anderen ist selbst bei Vorliegen der kompletten cDNA meist noch eine Umklonierung in einen geeigneten Expressionsvektor erforderlich. Der Zeitaufwand kann je nach Schwierigkeit des Projekts bei ein bis zwei Wochen, in Extremfällen aber durchaus auch bei mehreren Monaten bis Jahren liegen. Mit dem erfindungsgemäßen Verfahren kann ausgehend von einer bekannten Proteinsequenz ein Expressionskonstrukt mit einem für den gewünschten Organismus optimierten Codon Usage in ein bis zwei Tagen hergestellt werden. Hierfür muss der Organismus, in dem das Protein natürlicherweise exprimiert wird, gar nicht zur Verfügung



stehen, da die DNA-Sequenz aus der bekannten Proteinsequenz abgeleitet werden kann, ohne daß ein Template vorliegen muss. Mit verbesserten Proteinsequenzierungsmethoden wird es in Zukunft möglich sein, Proteine mit interessanten Enzymaktivitäten aus beliebigen Organismen zu sequenzieren und ohne den Umweg über die cDNA-Klonierung mittels des  
5 erfindungsgemäßen Verfahrens direkt in jedes gewünschte Expressionssystem zu überführen.

## 2. Herstellung von Designergenen und Designerproteinen

Eine weiterer Aspekt der vorliegenden Erfindung ist die einfache Herstellung von Designergenen bzw. Designerproteinen, d.h. die Kopplung funktioneller Domänen verschiedener Proteine, um  
10 beispielsweise Enzyme mit neuen oder veränderten Eigenschaften herzustellen. Bei Kenntnis der Röntgenkristallstruktur eines Proteins können dann sehr gezielte Veränderungen wie z.B. die Einfügung definierter Linkerdomänen oder ein Redesign einer Bindungstasche vorgenommen werden, um neue Funktionen oder veränderte Spezifitäten in Proteine einzuführen. Durch zielgerichtetes Proteindesign kann man beispielsweise regulierbare katalytische Zentren  
15 konstruieren, die durch eine Konformationsänderung des Proteins infolge der Bindung eines spezifischen Liganden aktiviert werden. Auf diese Weise können Designerproteine hergestellt werden, die z.B. auf die Bindung eines bestimmten Virusproteins hin eine Caspase-Aktivität entfalten, die dann in infizierten Zellen Apoptose auslöst. Erste Versionen solcher hochspezifischen Pharmaka sind bereits beschrieben worden; vgl. Vocero-Akbani A.M., Heyden  
20 N.V., Lissy N.A., Ratner L., Dowdy S.F., Nat Med, 1999 Jan, 5:1, 29-33. Weiterhin können Proteine dadurch stabilisiert werden, daß an bestimmten Positionen Aminosäuren eingebaut werden, die zusätzliche Salzbrücken bilden können. Somit kann die Toleranz gegenüber hohen Temperaturen verbessert werden, was unter anderem für die Waschmittelindustrie vorteilhaft ist. Sofern die Domänstrukturen bekannt sind, kann durch die punktgenaue Expression bestimmter  
25 funktioneller Regionen eine gewünschte enzymatische Aktivität von einer unerwünschten getrennt werden. Ebenso lassen sich Multienzymkomplexe konstruieren, die eine ganze Reihe verschiedener Reaktionen katalysieren können. Dadurch lassen sich Verbesserungen bei der Synthese vieler organischer Verbindungen erreichen oder manche Synthesen sogar erst ermöglichen. Dies eröffnet ganz neue Perspektiven, da viele organische Synthesen, bei denen  
30 heute noch umweltgefährdende Lösungsmittel und Katalysatoren eingesetzt werden müssen, in Zukunft durch solche maßgeschneiderten Biokatalysatoren ersetzt werden können.

### 3. Systematische Mutagenese als Ersatz für randomisierte Mutagenisierung

Eine häufig vorkommende Aufgabe in der biochemisch orientierten Molekularbiologie besteht darin, aus vielen Proteinvarianten diejenige herauszusuchen, die die höchste enzymatische Aktivität oder die stärkste Bindung an ein Substrat oder ein anderes Protein aufweist. Man geht  
5 dann meist so vor, daß man eine Reihe von zufälligen Mutationen von einer oder mehreren Aminosäuren einführt und die entstehenden Varianten in einem geeigneten Screening-Verfahren analysiert. Zwar ist es prinzipiell auch möglich, alle Mutanten separat herzustellen, jedoch wird dies aus Zeit- und Kostengründen selten durchgeführt. Bei einer randomisierten Mutagenisierung ist die Kontrolle über die entstehenden Mutanten naturgemäß sehr limitiert, zum einen da  
10 bestimmte Aminosäuresubstitutionen verfahrensbedingt häufiger gefunden werden als andere, zum anderen, da es sich kaum vermeiden lässt, daß bei diesem Verfahren auch Stopcodons eingeführt werden. Mit dem erfindungsgemäßen Verfahren lassen sich hingegen alle gewünschten Mutanten gezielt und ohne großen Aufwand darstellen und als Proteine exprimieren.

15

### 4. Herstellung synthetischer Gene, insbesondere zum Einsatz als DNA-Vakzine

In vielen Fällen ist es wünschenswert, die Proteinexpression bestimmter Gene in heterologen Systemen zu optimieren. Dies kann durch die Verwendung starker Promotoren sehr häufig nur zum Teil erreicht werden. Je nachdem, welcher Organismus für die Expression verwendet wird,  
20 kann sich die Verwendung bestimmter Codons für eine Aminosäure vorteilhaft oder nachteilhaft auf die erreichbare Genexpression auswirken. So sind beispielsweise viele retrovirale Genprodukte in eukaryotischen Zellen nur schlecht translatierbar, da sie meist sehr AT-reich sind und in höheren Eukaryonten seltene Codons benutzen. Speziell für den Einsatz solcher Gensequenzen als DNA-Vakzine ist es daher von großem Vorteil, wenn deren Codongebrauch  
25 für Säugerzellen optimiert ist. Desgleichen können bestimmte RNA-Strukturen zu einer Instabilität der Transkripte führen, was die Genexpression ebenfalls negativ beeinflussen kann. Solche Elemente können durch Codonveränderungen mit dem erfindungsgemäßen Verfahren ebenfalls leicht ausgeschaltet werden.

### 30 5. Analyse von Proteindomänen durch Deletions- oder Punktmutagenese

Die Analyse von Mutanten ist sehr oft das Mittel der Wahl bei der funktionellen Charakterisierung von Proteinen. Zwar existiert sowohl für die Herstellung von Deletions- wie auch

Punktmutanten eine Reihe etablierter Verfahren, jedoch sind diese in der Regel sehr zeitaufwendig und arbeitsintensiv. Deletionen werden meist durch Einführen von Linkersequenzen oder durch eine PCR mit Primern, deren Enden zu verschiedenen Teilsequenzen komplementär sind, hergestellt. Um eine ganze Serie definierter Deletionen zu erhalten, ist häufig ein  
5 Zweischrittverfahren notwendig, bei dem zunächst bestimmte Restriktionsschnittstellen eingeführt werden, über welche dann die gewünschten Deletionen eingeführt werden können. Mit entsprechend konzipierten Primern und einer Mehrfragmentligation lassen sich solche Deletionen zwar im Prinzip auch in einem Schritt herstellen, jedoch sind die Erfolgsaussichten dabei eher gering. In allen diesen Fällen muss die Wildtyp-DNA als Template vorliegen, was bei  
10 dem erfindungsgemäßen Verfahren nicht notwendig ist. Darüber hinaus können Deletionsmutanten hergestellt werden, da es gar nicht notwendig ist, Restriktionsschnittstellen einzuführen, für die man erst geeignete Stellen finden muss, damit die eingeführten Mutationen keine Veränderungen in der Proteinsequenz bewirken (sogenannte "Silent site" Mutationen). Auch die Herstellung von Doppel- oder Tripletmutanten ist mit dem erfindungsgemäßen Verfahren  
15 möglich. Für die funktionelle Kartierung eines Proteins können in dessen Gensequenz mittels der erwähnten "Silent site" Mutationen gleichzeitig auch Restriktionsschnittstellen für eine große Anzahl verschiedener Restriktionsendonukleasen eingeführt werden, mit deren Hilfe beliebige Deletionen hergestellt werden können. In vielen Fällen wird daher die klassische Mutationsanalyse verzichtbar und kann durch das schnellere und genauere erfindungsgemäße Verfahren  
20 ersetzt werden.

#### **6. Gekoppelte in vitro Transkriptions/Translationssysteme ("EasyPro™")**

Gekoppelte in vitro Transkriptions/Translationssysteme werden zur schnellen Synthese von Proteinen im analytischen Maßstab eingesetzt, z.B. für Bindungsstudien oder  
25 Kopräzipitationsassays. Hierfür werden die zu exprimierenden Sequenzen in einen Vektor kloniert, der einen Promoter für eine RNA-Polymerase enthält. Mit Hilfe dieser Polymerase wird mRNA transkribiert, die in einem RNA-depletierten Weizenkeim- oder Retikulozytenextrakt in das gewünschte Protein translatiert wird, das aufgrund der geringen Ausbeute und der einfacheren Nachweisbarkeit meist mit <sup>35</sup>S-Methionin oder Cystein radioaktiv markiert ist. Eine  
30 noch schnellere Alternative ist das auf dem erfindungsgemäßen Verfahren basierende EasyPro™-System. In einem Anchor-Oligonukleotid, welches einen T7 (SP6) Promoter, eine interne Ribosomenbindungsstelle sowie ein Hexahistidin-Tag enthält, wird durch Restriktion mit

XcmI ein einzelner Thymidinüberhang erzeugt, der direkt mit einem PCR-Produkt ligiert werden kann. Drei EasyPro™-Anchor-Oligonukleotide mit verschiedenen Leserastern sind ausreichend, um alle in der richtigen Orientierung ligierten PCR-Fragmente zu translatieren. Mittels terminaler Transferase oder durch Ligation eines entsprechenden Splinker-Oligonukleotids an das 3'-Ende des PCR-Produkts kann man zudem leicht einen künstlichen poly-A-Schwanz in das DNA-Template einführen, welcher das RNA-Transkript stabilisiert und dadurch für eine höhere Translationseffizienz sorgt. Ferner können die für das gewünschte Protein codierenden DNA-Sequenzen nach Spaltung mit einer Restriktionsendonuklease auch direkt an einen modifizierten EasyPro™-Anchor mit einem passenden 4 nt langen Überhang ligiert werden.

10

Ein weiterer Aspekt der vorliegenden Erfindung ist die Bereitstellung eines Minireaktors zur schnellen Synthese von Proteinen. In der unteren Reaktionskammer des Minireaktors findet die Transkription der an Streptavidin-beschichtete Beads gekoppelten Expressions-Anchor-Nukleinsäuresequenz statt. Die dabei entstehenden mRNAs werden über ihren 3'-poly-A-Schwanz an Oligo-dT gekoppelte Beads gebunden, welche sich ebenfalls in der unteren Reaktionskammer befinden. Dort läuft auch die Translation der mRNAs in einem Retikulozytenextrakt ab. Diese Kammer wird durch eine Ultrafiltrationsmembran mit einem MWCO (Molekulargewichtsausschluss) von ca. 200 kD von einer darüber liegenden zweiten Kammer abgetrennt. Diese enthält Beads, welche das Protein von Interesse binden können (z.B. Ni<sup>2+</sup>-NTA-Beads für Proteine mit einem Hexahistidin-Tag). Durch eine kontinuierliche Zufuhr von Pufferlösung mit frischen niedermolekularen Reaktanden (Aminoacyl-tRNAs, Ribonukleotidtriphosphate, CAP-Analogon und Creatinphosphat) wird die Produktion über längere Zeit hinweg aufrechterhalten. Gleichzeitig wird dadurch das synthetisierte Protein aus der unteren in die obere Reaktionskammer gedrückt, wo es an den Beads hängen bleibt.

Alternativ kann diese Kammer durch eine weitere Ultrafiltrationsmembran abgeschlossen werden, deren Ausschluß so gewählt ist, daß sie für Puffer und kleinere Moleküle, nicht aber für das gewünschte Protein permeabel ist. Dieses sammelt sich daher in der oberen Kammer an und kann von dort in aufgereinigter Form isoliert werden. Die dabei erzielbaren Ausbeuten sind nicht nur für die meisten analytischen Experimente ausreichend, sondern können sogar Proteinexpressionsexperimente im kleinen Maßstab ersetzen. Wenn es beispielsweise darum geht, die spezifische enzymatische Aktivität verschiedener Proteinmutanten zu bestimmen, mussten diese hierfür bislang in aufwendigen Vorversuchen kloniert, exprimiert und aufgereinigt

30



werden. Da praktisch alle diese Schritte in dem erfindungsgemäßen EasyPro™-Verfahren bereits integriert sind, wird dadurch ein erheblicher Zeitvorteil gegenüber konventionellen Methoden erreicht.

- 5 Mit einer Modifikation des vorstehend beschriebenen erfindungsgemäßen Verfahrens können einfach und kostengünstig Peptidbibliotheken hergestellt werden, welche unter anderem zum Epitopmapping von Antikörpern oder zur Identifizierung immunogener Epitope in Proteinen von Viren, Bakterien oder Pilzen benötigt werden (zur schnellen Etablierung serologischer Nachweissysteme). Hierfür werden modifizierte EasyPro™-Anchor-Oligonukleotide sukzessive
- 10 durch Splinkerligationen verlängert, so daß die für die gewünschten Peptide codierenden Sequenzen entstehen. Im letzten Schritt wird ein vorgefertigter Endsplinker anligiert, welcher für ein C-terminales Tag, ein Stopcodon sowie einen Poly-A-Schwanz codiert. Die Ligationsprodukte werden in dem beschriebenen Minireaktor transkribiert und translatiert. Die fertigen Peptide werden nach Abschluss der Translation und mehreren Waschschritten mit einer
- 15 spezifischen Protease, z.B. Enterokinase oder Faktor Xa, an der vom EasyPro™-Anchor-Oligonukleotid codierten Schnittstelle abgespalten und aus der oberen Reaktionskammer ausgewaschen. Mit Hilfe des C-terminalen Tags können diese an eine feste Phase für nachfolgende Tests gebunden werden. Die Peptide liegen zudem schon in aufgereinigter Form vor und können direkt für nachfolgende Applikationen eingesetzt werden. Da jeweils das gleiche
- 20 Anchor-Oligonukleotid verwendet wird und die benötigten Splinker-Oligonukleotide aus einem bereits vorgefertigten Teilstück in wenigen Schritten aufligiert werden können, sind die anfallenden Kosten geringer als bei einer konventionellen Peptidsynthese.

## **7. Herstellung von Ribozymen oder Aptameren**

- 25 Analog zu der oben beschriebenen Proteinsynthese lassen sich Anchor-Oligonukleotide mit einem T7 (SP6) Promoter auch zur Herstellung und Mutagenisierung von RNAs verwenden. Das System eignet sich insbesondere zur Synthese verschiedener Ribozyme, da die für die Ribozyme codierenden DNA-Sequenzen auf einem verlängerten Splinker-Oligonukleotid an ein Promotormodul auf einem Anchor-Oligonukleotid anligiert werden können. Vor allem können
- 30 genau definierte Ribozymtemplatebibliotheken erstellt werden, die sich per PCR leicht amplifizieren lassen. Mit dem erfindungsgemäßen Verfahren können Ribozymtemplatesequenzen auf das Nukleotid genau hergestellt werden, ohne daß hierfür



Klonierungsarbeiten notwendig sind. Durch Einführung von Linksequenzen kann man Ribozyme herstellen, die sich über eine zur Linksequenz komplementäre DNA/RNA an eine beliebige chemische Verbindung wie Peptide, Nukleinsäuren, aliphatische Kohlenwasserstoffe, Ester, Ether oder Alkohole koppeln lassen. Wenn diese Verbindung an eine feste Phase gebunden  
5 vorliegt, lassen sich Ribozyme selektionieren, die diese Bindung spalten. Diejenigen Ribozyme, die sich selbst von der Bindung an die feste Phase "befreit" haben, können durch reverse Transkription und nachfolgende asymmetrische PCR in einzelsträngige DNA-Moleküle überführt werden. Diese werden dann über die Linksequenz an ein entsprechend modifiziertes Anchor-Oligonukleotid hybridisiert und ligiert. Das verwendete Anchor-Oligonukleotid ist so  
10 konstruiert, daß es einen T7-Promoter enthält, über den mit Hilfe der T7-Polymerase wieder das Ribozym erhalten werden kann. Durch die Verwendung einer ungenauen Reversen Transkriptase (z.B. HIV RT) lassen sich zufällige Mutationen einführen. Der Selektionsdruck kann durch immer kürzere Inkubationen erhöht werden, so daß präferentiell Ribozyme mit einer hohen Aktivität amplifiziert werden. Analog lassen sich nach demselben Prinzip auch Ribozyme mit der  
15 Fähigkeit selektionieren, eine Bindung zur festen Phase zu vermitteln.

#### **8. Verwendung von mit dem erfindungsgemäßen Verfahren produzierten ssDNAs in der Diagnostik (PathoCheck™)**

Bei der Diagnostik von Infektionskrankheiten wird üblicherweise häufig ein direkter  
20 Erregernachweis z.B. durch PCR verlangt. Besonders in der Transfusionsmedizin ist es wichtig, kontaminierte Blutproben sicher zu erkennen und auszusortieren. Die hierfür üblicherweise eingesetzten serologischen Assays können dies nur dann gewährleisten, wenn die Infektion des Spenders bereits einige Zeit zurückliegt, so daß bereits Antikörper gebildet worden sind. Während eines Fensters von bis zu 12 Wochen (in Extremfällen auch länger) sind beispielsweise  
25 bei der HIV-Infektion noch keine Antikörper im Blut nachweisbar, obwohl bereits eine massive Virusreplikation stattfindet. Da eine routinemäßige PCR-Untersuchung aller Proben aus Kostengründen vielerorts kaum machbar ist, wird diese (wenn überhaupt) an Pools von Einzelspenden durchgeführt. Die Problematik dabei ist, daß dadurch die an sich sehr hohe Sensitivität sinkt, da die Menge des für die Analyse eingesetzten Materials nicht beliebig erhöht  
30 werden kann. Bei Viruserkrankungen wie HIV, bei denen ein Großteil der Viren extrazellulär vorliegt, kann dies durch eine Konzentrierung der Viren durch Zentrifugation noch einigermaßen ausgeglichen werden, bei vorwiegend zellassozierten Viren funktioniert das allerdings meist

weniger gut. Man kann zwar zunächst DNA oder RNA aus den Blutzellen isolieren, jedoch nur einen Bruchteil davon in der PCR-Reaktion einsetzen, da sonst unspezifische PCR-Produkte überhandnehmen. Daher muss in solchen Fällen eine Vorselektion des zu amplifizierenden Materials durchgeführt werden. Hierfür wird ein einzelsträngiges, mit dem erfindungsgemäßen Verfahren hergestelltes Produkt eingesetzt, das mit einem modifizierten Anchor-Oligonukleotid hergestellt wird. In diesem Fall verwendet man ein Anchor-Oligonukleotid aus zwei getrennten komplementären Strängen, von denen der eine am 5'-Ende modifiziert, z.B. biotinyliert, ist, der andere am 3'-Ende blockiert ist. Nach der Synthese der Virus-Sequenz wird der nicht-biotinylierte Strang durch Waschen mit einer denaturierenden Lösung abgetrennt, so daß eine einzelsträngige Antisense-DNA verbleibt. Diese kann mit dem 5'-biotinylierten Teil-Ancor-Oligonukleotid und einem endständigen Oligonukleotid amplifiziert werden, sofern mehr Material benötigt wird. Von diesem PCR-Produkt ist nur ein Strang biotinyliert, der andere kann durch Denaturierung abgetrennt werden. Diese Antisense-DNA kann nun dafür eingesetzt werden, um virale RNAs oder DNAs aus einem komplexen Gemisch wie einem Zelllysate oder einer Nukleinsäurepräparation anzureichern, indem man diese miteinander hybridisiert, die Hybride an einen Streptavidin-beschichteten Träger (Support) bindet und nicht-hybridisierte Komponenten unter stringenten Bedingungen wegwäscht. In einem zweiten Schritt können die angereicherten RNAs oder DNAs dann mit einer konventionellen PCR mit Primern aus dem nicht-hybridisierten Teil der RNA oder DNA amplifiziert und nachgewiesen werden. Dies kann üblicherweise über Gelelektrophorese der Produkte geschehen oder durch Fluoreszenzanalyse oder durch einen nachgeschalteten ELISA, unter der Voraussetzung, daß ein entsprechend modifizierter Primer verwendet wurde. Vorteile des erfindungsgemäßen Verfahrens sind, daß fast beliebig hohe Mengen an Ausgangsmaterial eingesetzt werden können, was die Sensitivität der Analyse verbessert, daß auch mehrere Targets gleichzeitig untersucht und bei Verwendung verschieden fluoreszenzmarkierter Primer auch differenziert werden können und daß es für beliebige Pathogene wie Bakterien, Pilze oder Viren einsetzbar ist. Durch die Voranreicherung der zu amplifizierenden Sequenzen werden nebenbei auch Hintergrundprobleme deutlich reduziert. Bei entsprechender Miniaturisierung kann eine große Anzahl verschiedener Pathogene gleichzeitig auf einem Chip getestet werden, wodurch die Analysekosten extrem gesenkt werden können.

## 9. "Gene Profiling" (GPro™)

In der molekularbiologischen Forschung und zunehmend auch in der molekularen Diagnostik wird die Expression bestimmter Gene auf der RNA-Ebene quantitativ untersucht. Standardwerkzeuge hierfür sind der Northern-Blot, das S1-Mapping oder der "Ribonuclease Protection Assay" (RPA), meist in Verbindung mit radioaktiv markierten Sonden. Die vorstehend  
5 beschriebenen einzelsträngigen DNAs können auch für diese Aufgabe verwendet werden. Eine vorhergehende Aufreinigung der zu analysierenden RNAs, die meist eine zusätzliche Fehlerquelle darstellt, ist hierzu nicht notwendig. Ähnlich wie bei dem erfindungsgemäßen PathoCheck™-Verfahren werden die zu untersuchenden mRNAs mit einem Überschuss eines modifizierten z.B. biotinylierten Anchor-Oligonukleotids mit genspezifischen einzelsträngigen  
10 Antisense-DNAs hybridisiert und an einer z.B. Streptavidin-beschichteten festen Phase immobilisiert. Nach dem Auswaschen aller Proteine, nicht relevanter Nukleinsäuren und sonstiger Verunreinigungen werden die Ziel-mRNAs mit einer Reihe von direkt oder indirekt fluoreszenzmarkierten Splinkersequenzen, welche zu einem anderen Teil dieser mRNAs komplementär sind, nachgewiesen. Durch die Verwendung verschiedener genspezifischer  
15 Antisense-DNAs und unterschiedlich markierter Nachweissplinker-Oligonukleotide kann die Expression mehrerer Gene gleichzeitig analysiert werden. Das ganze Verfahren lässt sich ohne großen Aufwand vollständig automatisieren. Sofern die zu analysierenden Gewebe nicht übermäßig viel RNase enthalten, reicht eine Lyse in chaotropen Puffern und/oder Zugabe von RNasin aus, die Integrität der RNAs zu gewährleisten. Wenn maximale Sensitivität wichtiger ist  
20 als der gleichzeitige Nachweis verschiedener mRNAs in einem Reaktionsansatz, können anstelle der fluoreszenzbasierten Nachweisreagenzien auch Antikörper eingesetzt werden, welche an ein polyvalentes Sekundärreagenz binden wie ein anti-Maus-Ig-Peroxidase-Polymer. Diese Komplexe werden dann in einer nachgeschalteten Enzymreaktion detektiert z.B. durch die beim Umsetzen eines geeigneten Substrats entstehende Chemilumineszenz. Für besonders häufig  
25 gemachte Untersuchungen können entsprechende GPro™-Kits mit synthetischen Kontroll-mRNAs als quantitative Standards bereits fertig konfektioniert werden.

#### **10. Allelidentifizierung durch hybridvermittelte Ligation (LIMA™; Ligation mediated Identification of Mutant Alleles)**

30 Besonders in der Pränataldiagnostik von Erbkrankheiten, aber auch zur Bestimmung der individuellen Sensitivität gegenüber verschiedenen Arzneimitteln muss der Genotyp bestimmter Allele festgestellt werden. In der Regel geschieht dies durch eine PCR-Amplifikation des zu

untersuchenden Lokus aus der genomischen DNA und anschließende Restriktionsanalyse oder Sequenzierung. Im ersten Fall bedingt dies eine gelelektrophoretische Auftrennung der Restriktionsfragmente, die nicht ohne weiteres automatisierbar ist. Das trifft auch im zweiten Fall zu, sofern nicht mit der Chipsequenzierungsmethode gearbeitet wird, die jedoch noch nicht  
5 ausgereift ist. Auch für diesen Aspekt lassen sich erfindungsgemäß hergestellte DNA-Fragmente verwenden. Voraussetzung hierfür ist, daß es sich um bekannte, molekular identifizierte Allele handeln muss. Ein erfindungsgemäß hergestelltes Anchor-Oligonukleotid wird dann so konstruiert, daß es an eine Genregion hybridisiert, die unmittelbar vor der Mutation liegt. Ein weiteres Oligonukleotid, das ein oder mehrere fluoreszierende Markierungen enthält, hybridisiert  
10 an die direkt angrenzende 3'-Region des Gens, so daß die beiden freien Enden der erfindungsgemäß hergestellten Anchor-Oligonukleotid-DNA und des fluoreszenzmarkierten Oligonukleotids bei durchgängiger Hybridbildung direkt nebeneinander zu liegen kommen und miteinander ligiert werden können. Sofern die Sequenz an dieser Stelle abweicht, kommt es nicht zur Anlagerung der Enden und damit auch nicht zur Ligation. Stattdessen kann z.B. ein mit einer  
15 anderen Fluoreszenz markiertes Oligonukleotid an die entsprechende mutierte Sequenz binden, wodurch eine andere Markierung an den biotinylierten Anchor ligiert wird. Durch Laseranregung werden die jeweils gebundenen Fluoreszenzfarbstoffe und damit die jeweiligen Allele identifiziert. Zur Erhöhung der Sensitivität des erfindungsgemäßen Verfahrens kann auch in diesem Fall eine asymmetrische PCR vorgeschaltet werden, welche den zu untersuchenden  
20 Lokus amplifiziert. Bei einheitlichen Reaktionsbedingungen für die PCR und Hybridisierung ist es möglich, mehrere verschiedene Allele gleichzeitig aus einer Probe zu bestimmen.

### **11. Direkte Interaktionsanalyse von Protein Arrays (LISPA™)**

Mit dem Erfolg des "Human Genome Project" steht als eine der nächsten Aufgaben die  
25 Klassifizierung der ca. 50.000 menschlichen Gene an. Nicht nur in der Grundlagenforschung, sondern auch in dem sich rasch fortentwickelnden Gebiet der molekularen Medizin ist es wichtig zu verstehen, was diese Gene tun, wie sie in welchen Situationen miteinander kooperieren, welche Proteine, Peptide oder niedermolekularen Stoffe an welche anderen Proteine binden etc. Ein erster Anhaltspunkt für solche Kooperationen zwischen Proteinen ist meist ein direkter  
30 physischer Kontakt der jeweiligen Genprodukte. Um solche Bindungen in vitro auf der Proteinebene untersuchen zu können, benötigt man in der Regel aufgereinigte Proteinpräparationen. Bei 50.000 Proteinen ist dies jedoch nicht ohne weiteres möglich. Man behilft



sich daher zumeist mit genetischen Methoden wie z.B. dem sogenannten "Yeast Two Hybrid Screen", um mögliche Interaktionspartner zu identifizieren. So erfolgreich diese Methode bisher auch eingesetzt worden ist, ist sie dennoch extrem artefaktanfällig, umständlich und für die Komplexität der nun anstehenden Aufgabe ungeeignet. Diese Aufgabe kann mit einer Kombination des erfindungsgemäßen Verfahrens, des Sloning<sup>TM</sup>-Verfahrens, und dem erfindungsgemäßen EasyPro<sup>TM</sup>-Verfahren in Verbindung mit einem Biochip durchgeführt werden. Mittels einer Automatisierung können die kompletten 50.000 Gene synthetisiert, exprimiert und mit einem geeigneten Tag zur Immobilisierung in Reaktionskammern eines Biochips versehen werden. Für Bindungsstudien mit einem fluoreszenzmarkierten Protein oder einer niedermolekularen chemischen Verbindung ist eine Menge von  $10^7$  bis  $10^8$  Molekülen üblicherweise ausreichend. In Kavitäten von  $100 \times 200 \times 60 \mu\text{m}$  können etwa 1 Nanoliter einer Proteinlösung deponiert werden, dies entspricht bei einem 100 kD Protein und einer Konzentration von 5 mg/ml ca.  $3 \times 10^{10}$  Molekülen. Geht man davon aus, daß die tatsächliche Bindungskapazität pro Kavität bei ca. 1% dieses Werts liegt, ist auch bei relativ großen Proteinen noch ausreichend Material vorhanden. Wenn die einzelnen Kavitäten ca.  $30 \mu\text{m}$  auseinanderliegen, so könnte die gesamte Bibliothek von 50.000 Proteinen auf einem Chip von nur  $20 \text{ cm}^2$  untergebracht werden. Ein Laser mißt die Fluoreszenz in sämtlichen Kavitäten vor und nach der Bindung des fluoreszenzmarkierten Zielmoleküls, woraus man die Stärke der Interaktion berechnet. Kavitäten, in denen nur das Tag präsentiert wird, dienen als unspezifische Kontrolle. Mit Hilfe solcher Proteinarrays lassen sich beispielsweise bislang nicht feststellbare Bindungen von Arzneimitteln an zelluläre Proteine detektieren oder auch komplizierte Signaltransduktionskaskaden nachvollziehen.

## 12. Parallele Analyse von mRNAs mit immobilisierten Nukleinsäure Arrays (PAMINA<sup>TM</sup>)

Einer der Schwerpunkte der modernen Arzneimittelforschung besteht in dem gezielten Eingriff in die Expression einzelner Gene. Dazu muss der Einfluss von neuen Wirkstoffen auf die Expression anderer Gene möglichst umfassend untersucht werden. Bei Signalübertragungsprozessen, der Zelldifferenzierung oder bei krankheitsinduzierten metabolischen Veränderungen wird oft eine ganze Kaskade verschiedener Gene an- oder abgeschaltet. Aufgrund der Komplexität der Genexpression in höheren Organismen ist es aber bis jetzt praktisch nicht möglich, mehr als eine Handvoll Gene gleichzeitig zu analysieren. Mit der



Sequenzierung des menschlichen Genoms werden in Zukunft jedoch die Grundvoraussetzungen für eine umfassende parallele Analyse der gesamten Genexpression einer Zelle geschaffen werden. Aus den vorliegenden Sequenzinformationen lassen sich durch computerisierte Sequenzvergleiche zunächst die Regionen in den einzelnen Genen identifizieren, die die geringste Homologie untereinander aufweisen, also den höchsten Grad an Spezifität für das jeweilige Gen. Aus diesen Genabschnitten können genspezifische, einzelsträngige Antisense-Anchor-DNAs abgeleitet werden, die in einem Array auf einem Biochip immobilisiert werden. Die Antisense-Anchor-DNAs können so konzipiert werden, daß die Schmelztemperaturen sämtlicher RNA/DNA-Hybride in einem engen Fenster liegen. Durch Hybridisierung der gesamten RNA, bzw. der polyA<sup>+</sup>-RNA der zu untersuchenden Zellen an diesen Array unter Bedingungen maximaler Stringenz werden in jeder Kavität des Biochips RNA/DNA-Hybride erzeugt, falls die korrespondierende mRNA exprimiert wird. In einem 2. Schritt werden die immobilisierten Antisense-Anchor-DNAs an dem hybridisierten RNA-Template durch Behandlung mit einer RNaseH Reversen Transkriptase und modifizierten Nukleotiden, die vorzugsweise fluoreszenzmarkiert sind, verlängert. Nach mehreren Waschvorgängen, um nicht inkorporierte Nukleotide abzutrennen, können die cDNA-Reaktionsprodukte analog zu der beschriebenen LISPA<sup>TM</sup>-Technik durch Laserabtastung der einzelnen Kavitäten gemessen werden.

## 20 Ausführungsbeispiele

### 1. Herstellung der Anchor- und Splinkeroligonukleotide

Die Anchor- und Splinker-Oligonukleotide wurden nach dem von Sinha N.D., Biernat J., McManus J., Köster H., *Nucleic Acids Res*, 1984 Jun, 12:11, 4539-57 beschriebenen Standardverfahren hergestellt oder durch eine Synthese in großem Maßstab, die nacheinander geviertelt wurde oder durch Simultansynthese auf Cellulosemembranen.

### 2. Markierung mit einer Modifikation

Modifikationen in den Oligonukleotiden wurden mit Standardverfahren durchgeführt.

## 30 3. Kopplung an die Matrix

Zu 10 µl Streptavidin-gekoppelten magnetisierten Beads (MERCK) in einem Gesamtvolumen von 50 µl in 1xTE/1 M NaCl, pH 7.5 wurden 20-200 pmol Biotin-markierte kinasierte Anchor-

Oligonukleotide zugesetzt und 30 min bei Raumtemperatur auf einem Roller inkubiert. Anschließend wurden nicht gebundene Anchor-Oligonukleotide durch einen dreimaligen Pufferwechsel von je 500 µl 1xTE, pH 7.5 gewaschen.

### 5 3. Erster Ligationsschritt

Die Ligation erfolgte bei 4°C, 16°C, Raumtemperatur bzw. 37°C (Standard 16°C) in einem Volumen von 50 µl in 1 x Ligasepuffer (Boehringer Mannheim) mit 1 bis 5 Einheiten T4 DNA Ligase (Boehringer Mannheim oder New England Biolabs) für 15 bis 60 Minuten. Für die Ligation wurden in der Regel 20 pmol phosphoryliertes Anchor-Oligonukleotid eingesetzt. Am  
10 5'-Ende phosphorylierte Splinker-Oligonukleotide wurden in 1,5 bis 5fachem molaren Überschuß zugegeben. Nach der Reaktion wurden Ligase und nicht-ligierte Splinker-Oligonukleotide durch drei Pufferwechsel von je 500 µl 1xTE, pH 7.5 gewaschen. Danach wurde zu den gewaschenen Beads 40 µl eines Restriktionsmixes gegeben, der das Splinker-spezifische Restriktionsenzym Eco31I in 1,25x Restriktionspuffer (Puffer A von Boehringer  
15 Mannheim bzw. Puffer 4 von New England Biolabs) enthielt. Danach wurde wie vorstehend beschrieben gewaschen.

### 5. Zweiter Ligationsschritt

Es wurden vier weitere Ligationen mit weiteren Splinker-Oligonukleotiden nach der unter Punkt  
20 4 aufgeführten Vorschrift durchgeführt.

### 6. Transposition

Nach der 5. Ligation wurde nach dem Waschen ein Mix des Anchor-spezifischen Restriktionsenzyms Esp3I bzw. BpiI in den entsprechenden herstellerspezifischen Puffern  
25 hinzugefügt und 30 bis 60 Minuten bei 37°C inkubiert. Nach der Reaktion wurde der komplette Mix mitsamt den abgeschnittenen aufligierten Splinker-Oligonukleotiden entfernt, in einem separaten Reaktionsgefäß 15 Minuten bei 65°C hitzebehandelt, um das Restriktionsenzym zu inaktivieren und sodann in ein weiteres Reaktionsgefäß mit entsprechend aufligierten, an magnetisierte Streptavidin-Beads gekoppelte Anchor-Oligonukleotide ligiert.

30

### 7. Restriktionskontrolle ligierter Fragmente

Zur Kontrolle der korrekten Größe der geschnittenen Splinker-Oligonukleotide wurde ein 5 µl

Aliquot der Reaktion auf einem 18%igen 1xTBE-Polyacrylamidgel aufgetrennt, mit 0,01% SYBR-Gold™ in 1 x TBE für 10 min angefärbt und mit UV-Licht sichtbar gemacht. Auf einem solchen Gel können Längenunterschiede von 1-2 Basen erkannt werden.

5

## Patentansprüche

1. Verfahren zur Herstellung eines Nukleinsäuremoleküls, umfassend die Schritte
  - a) Kopplung eines Oligonukleotids mit einem Ende an eine feste Matrix, wobei die  
10 Kopplung über eine Modifikation erfolgt, und das Oligonukleotid eine Erkennungssequenz für ein TypIIS Restriktionsenzym enthält, welches außerhalb seiner Erkennungssequenz schneidet,
  - b) Zugabe eines weiteren, mindestens teilweise doppelsträngigen Oligonukleotids mit einer anderen Erkennungssequenz für ein TypIIS Restriktionsenzym, welches  
15 außerhalb seiner Erkennungssequenz schneidet, als in Schritt a), wobei dieses Oligonukleotid nicht an die Matrix binden kann
  - d) Ligation der Oligonukleotide aus Schritt a) und b) in der durch die Blockierung der nicht zu ligierenden Enden fest gelegten Orientierung
  - e) Entfernen nicht verbrauchter Reaktanden sowie Enzyme
  - 20 e) Spaltung des Ligationsprodukts aus Schritt c) mit einem TypIIS Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung im Oligonukleotid aus Schritt a) stattfindet,
  - f) Abtrennen des so verlängerten Nukleinsäuremoleküls vom Reaktionsgemisch.
- 25 2. Verfahren zur Herstellung eines Nukleinsäuremoleküls, umfassend die Schritte
  - a) bis d) gemäß Anspruch 1,
  - e) Spaltung des Ligationsprodukts aus Schritt c) mit einem TypIIS Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung in der  
Nukleinsäuresequenz des Oligonukleotids aus Schritt b) stattfindet,
  - 30 f) Abtrennen des Reaktionsgemisches vom in Schritt e) erhaltenen verlängerten Oligonukleotid aus Schritt a),
  - g) mindestens einmaliges Wiederholen der Schritte b) bis f).

3. Verfahren nach Anspruch 2, ferner umfassend
- h) Spaltung des erhaltenen Nukleinsäuremoleküls mit einem TypII Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung im Oligonukleotid aus Schritt a) stattfindet, und gegebenenfalls
  - 5 i) Spaltung des erhaltenen Nukleinsäuremoleküls mit einem TypII Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, wobei die Spaltung in der Nukleinsäuresequenz des Oligonukleotids aus Schritt b) stattfindet.
4. Verfahren nach Anspruch 3, ferner umfassend Abtrennen des erhaltenen Nukleinsäure-
- 10 moleküls vom Reaktionsgemisch.
5. Verfahren nach einem der Ansprüche 1 bis 4, wobei das in Schritt b) eingesetzte Oligonukleotid ein durch das Verfahren nach Anspruch 1 bis 4 hergestelltes Nukleinsäuremolekül ist.
- 15
6. Verfahren nach einem der Ansprüche 1 bis 5, wobei nach Schritt c) als Schritt c)' eine Exonuklease- und/oder Phosphatase-Reaktion durchgeführt wird.
7. Verfahren nach Anspruch 6, wobei das Reaktionsgemisch des Schrittes c)' nach der
- 20 Reaktion entfernt wird.
8. Verfahren nach einem der Ansprüche 1 bis 7, wobei das Oligonukleotid aus Schritt a) an dem Ende, welches nicht an die Matrix gekoppelt ist, einen Teil einer Erkennungssequenz für ein TypII Restriktionsenzym enthält, welches außerhalb seiner Erkennungssequenz schneidet, und
- 25 der andere Teil der Erkennungssequenz für dieses Restriktionsenzym vom Oligonukleotid aus Schritt b) stammt.
9. Verfahren nach Anspruch 1 bis 8, wobei die Modifikation ein Biotinrest, ein Digoxigeninrest, ein Fluoresceinisothiocyanatrest, eine Aminoverbindung oder ein Succinylester ist.
- 30
10. Verfahren nach einem der Ansprüche 1 bis 9, wobei das Oligonukleotid aus Schritt a) und/oder b) über einen Loop verfügt.



11. Verfahren nach Anspruch 10, wobei das Oligonukleotid aus Schritt a) über eine Modifikation im Loopbereich an die feste Matrix gekoppelt ist.
- 5 12. Verfahren nach einem der Ansprüche 1 bis 11, wobei die feste Matrix ein Kügelchen (bead), vorzugsweise aus Glas oder Polystyrol, ein Objektträger, ein DNA Chip, die Vertiefung einer Mikrotiterplatte (well) oder ein Reaktionsröhrchen ist.
13. Verfahren nach einem der Ansprüche 1 bis 12, wobei die feste Matrix einen Streptavidin-  
10 rest, einen anti-Digoxigenin-Antikörper oder anti-Fluoresceinisothiocyanat-Antikörper umfasst.
14. Verfahren nach einem der Ansprüche 1 bis 13, wobei die Oligonukleotide aus Schritt a) und b) an ihren zu ligierenden Enden über zueinander komplementäre Einzelstrangüberhänge verfügen.
- 15 15. Verfahren nach Anspruch 14, wobei die Einzelstrangüberhänge 1, 2, 3, 4 oder 5 Nukleotide lang sind.
16. Verfahren nach Anspruch 1 bis 15, wobei die verschiedenen TypIIS  
20 Restriktionsendonukleasen durch analog schneidende Ribozyme ersetzt werden
17. Verfahren nach einem der Ansprüche 1 bis 16, wobei das Oligonukleotid in Schritt b) ein PCR-Produkt, ein Plasmidvektor, eine Phagen- oder Virus-DNA, ein künstliches Chromosom oder eine weitere synthetische DNA ist.
- 25 18. Kit zur Herstellung einer Nukleinsäuresequenz nach dem Verfahren nach einem der Ansprüche 1 bis 17, umfassend
- a) eine Bibliothek von 1 bis 1.048.576 verschiedenen Oligonukleotiden enthält, wobei die Oligonukleotide über eine Modifikation an einem Ende an eine feste Matrix  
30 koppelbar sind und das Oligonukleotid eine Erkennungssequenz oder einen Teil der Erkennungssequenz für ein TypIIS Restriktionsenzym enthält, welches außerhalb seiner Erkennungssequenz schneidet,

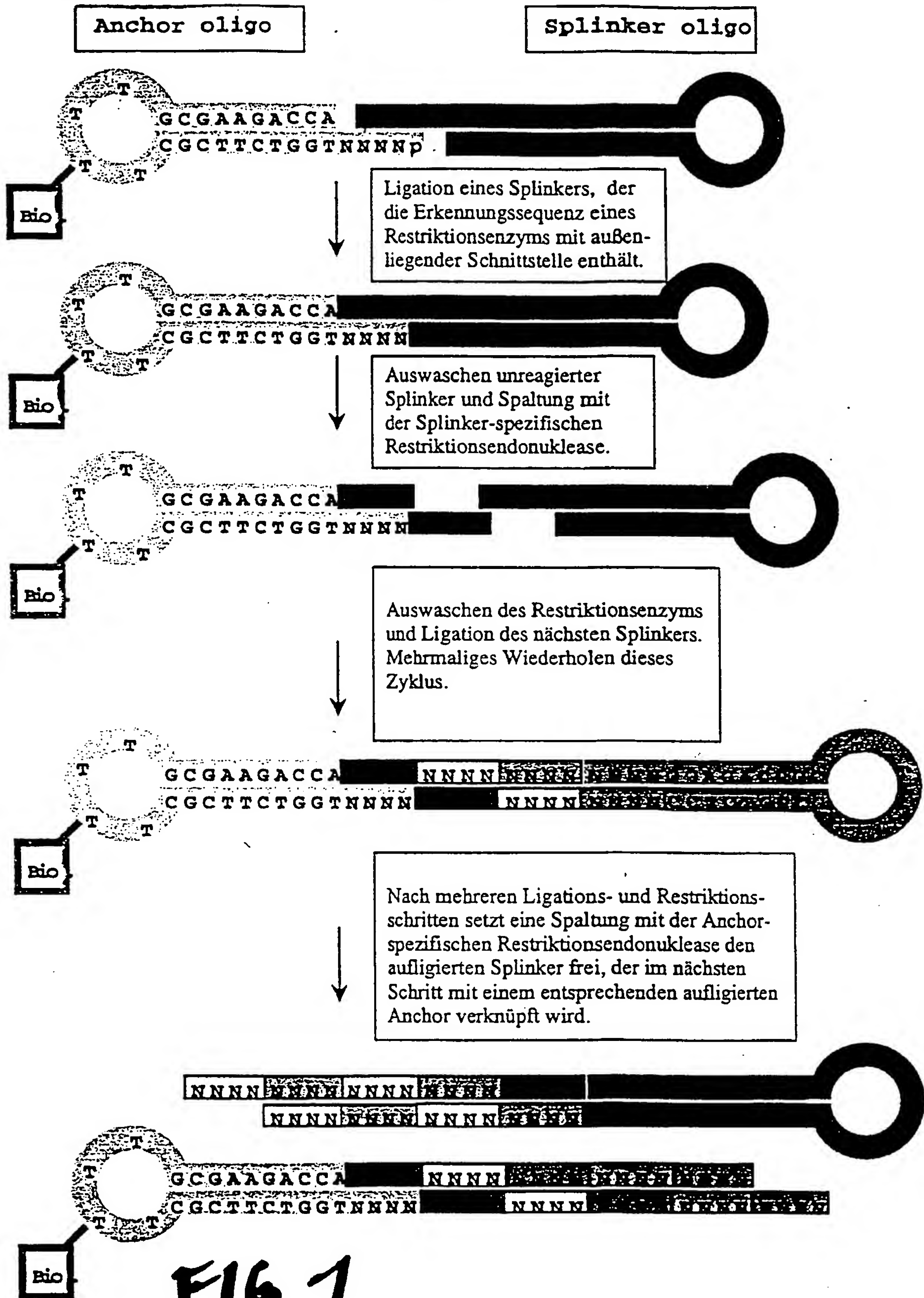
- 5           b) eine weitere Bibliothek von 4 bis 1.048.576 verschiedenen Oligonukleotiden enthält, wobei jedes der Oligonukleotide eine Erkennungssequenz für ein TypIIS Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, enthält, dass von dem TypIIS Restriktionsenzym aus a) verschieden ist, und gegebenenfalls den anderen Teil der Erkennungssequenz des Restriktionsenzym aus Schritt a) enthält,
- c) eine feste Matrix,
- d) Reservoirre der für die Herstellung des Nukleinsäuremoleküls benötigten Enzyme und/oder anderer Reagenzien.

10

19. Vorrichtung zur automatisierten Herstellung eines Nukleinsäuremoleküls nach dem Verfahren nach einem der Ansprüche 1 bis 16, dadurch charakterisiert, dass sie

- 15           a) eine Bibliothek von 1 bis 1.048.576 verschiedenen Oligonukleotiden enthält, wobei die Oligonukleotide über eine Modifikation an einem Ende an eine feste Matrix koppelbar sind und das Oligonukleotid eine Erkennungssequenz oder einen Teil der Erkennungssequenz für ein TypIIS Restriktionsenzym enthält, welches außerhalb seiner Erkennungssequenz schneidet,
- 20           b) eine weitere Bibliothek von 4 bis 1.048.576 verschiedenen Oligonukleotiden enthält, wobei jedes der Oligonukleotide eine Erkennungssequenz für ein TypIIS Restriktionsenzym, welches außerhalb seiner Erkennungssequenz schneidet, enthält, dass von dem TypIIS Restriktionsenzym aus a) verschieden ist, und gegebenenfalls den anderen Teil der Erkennungssequenz des Restriktionsenzym aus Schritt a) enthält,
- c) eine feste Matrix,
- 25           d) Reservoirre der für die Herstellung des Nukleinsäuremoleküls benötigten Enzyme und/oder anderer Reagenzien, und
- e) ein Steuerungsprogramm, das einzelne Oligonukleotide aus a) und b) identifizieren, mit der festen Matrix aus c) und den benötigten Enzymen und/oder anderen Reagenzien aus d) zusammenbringen und die Abfolge von Herstellungsschritten
- 30           bestimmen und durchführen kann.

20. Verwendung des nach einem der vorstehenden Verfahren hergestellten Nukleinsäuremoleküls als DNA-Vakzine, zur Analyse von Proteindomänen, als Matrize für Designerproteine, zur schnellen Proteinsynthese, zur Herstellung von Ribozymen oder Aptameren, als Sonde zum Nachweis pathogener Mikroorganismen, als Sonde zum Nachweis der Expression von Genen,
- 5 zum Nachweis allelspezifischer Mutationen, zum Nachweis von Protein/Protein-Bindung, Protein/Peptid-Bindung und/oder der Bindung niedermolekularer Stoffe an Proteine.



2/8

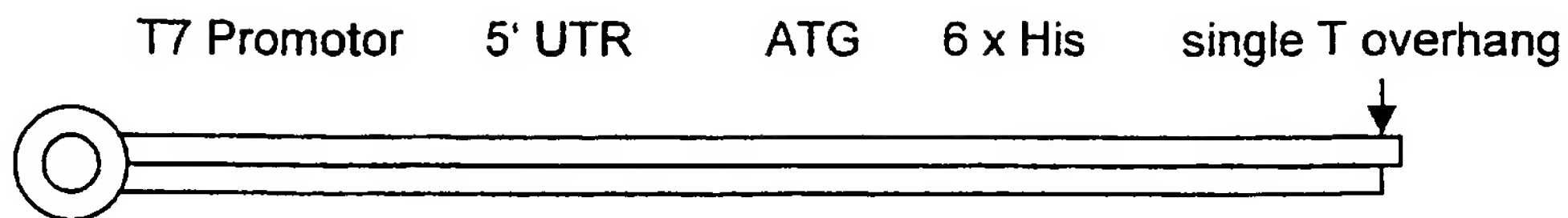


Fig. 2

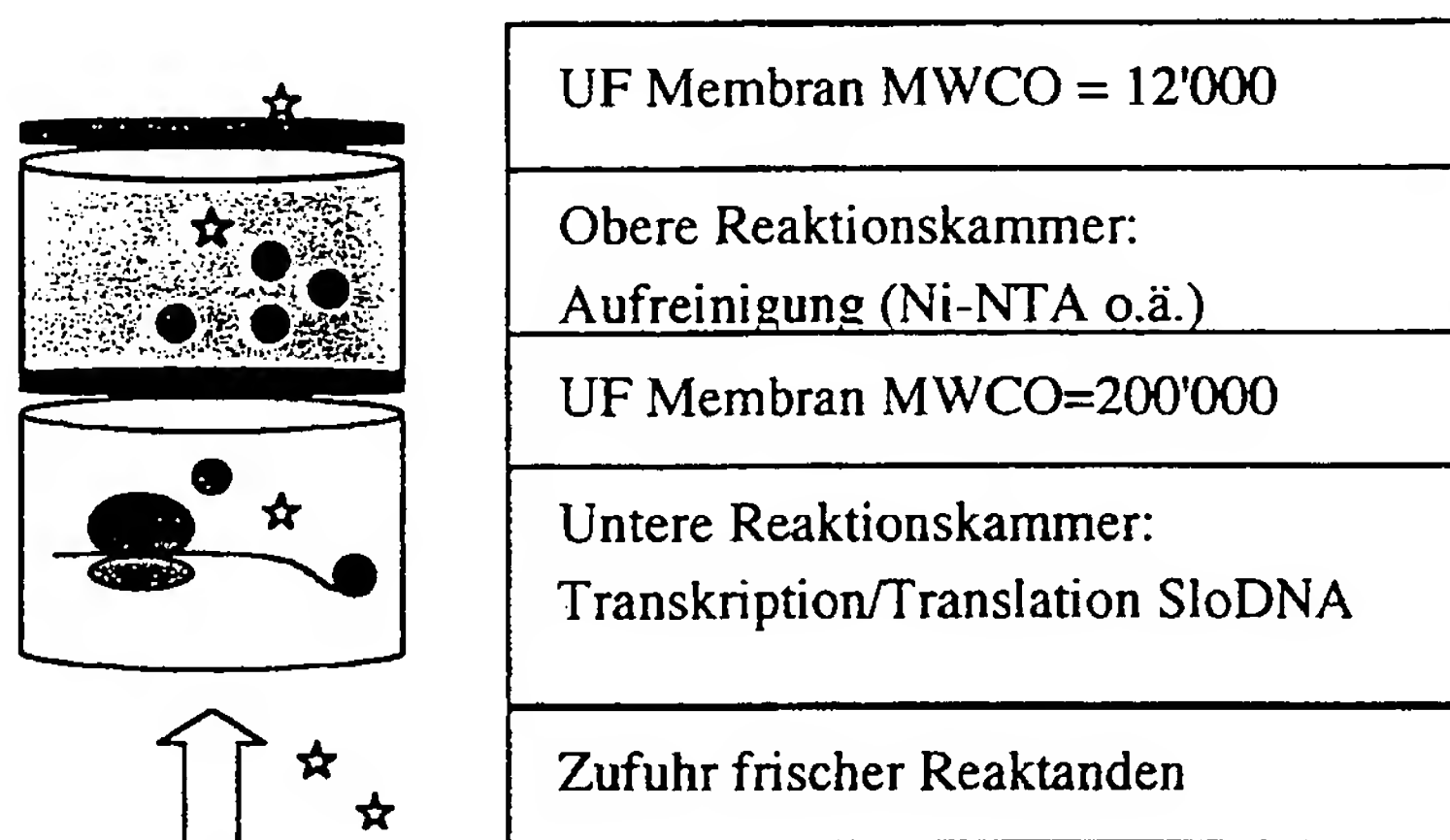


Fig. 3

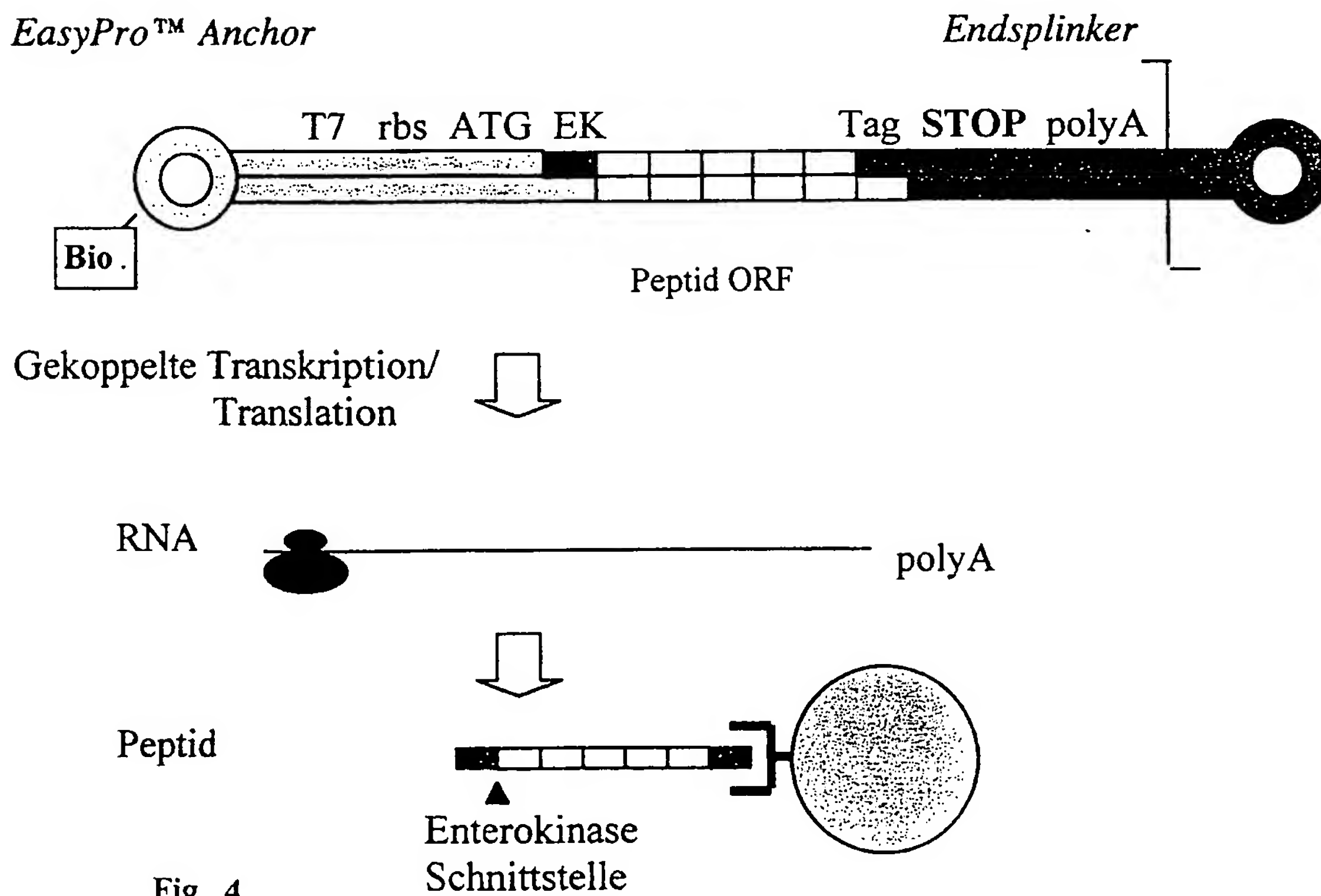


Fig. 4



3/8

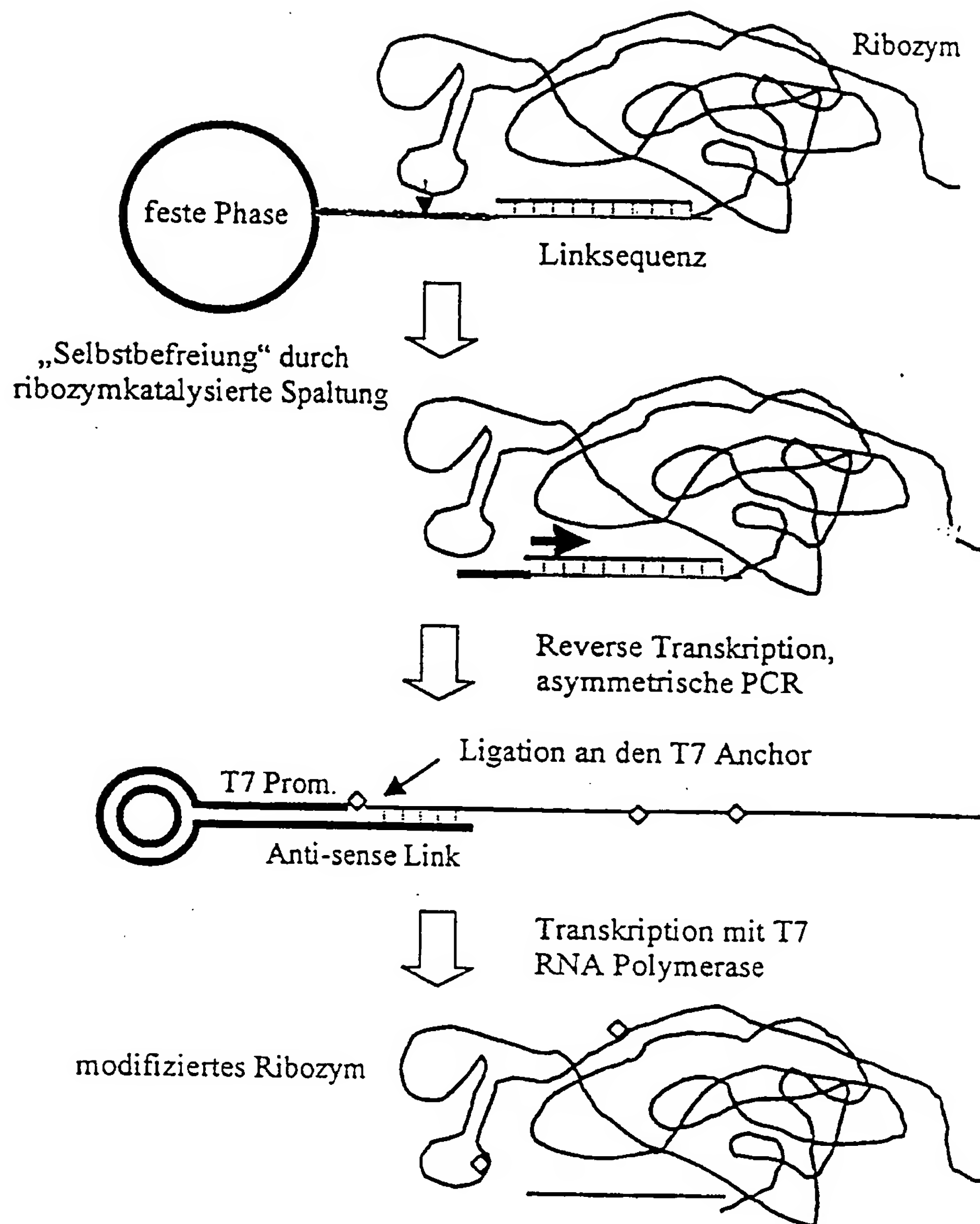


Fig. 5

4/8

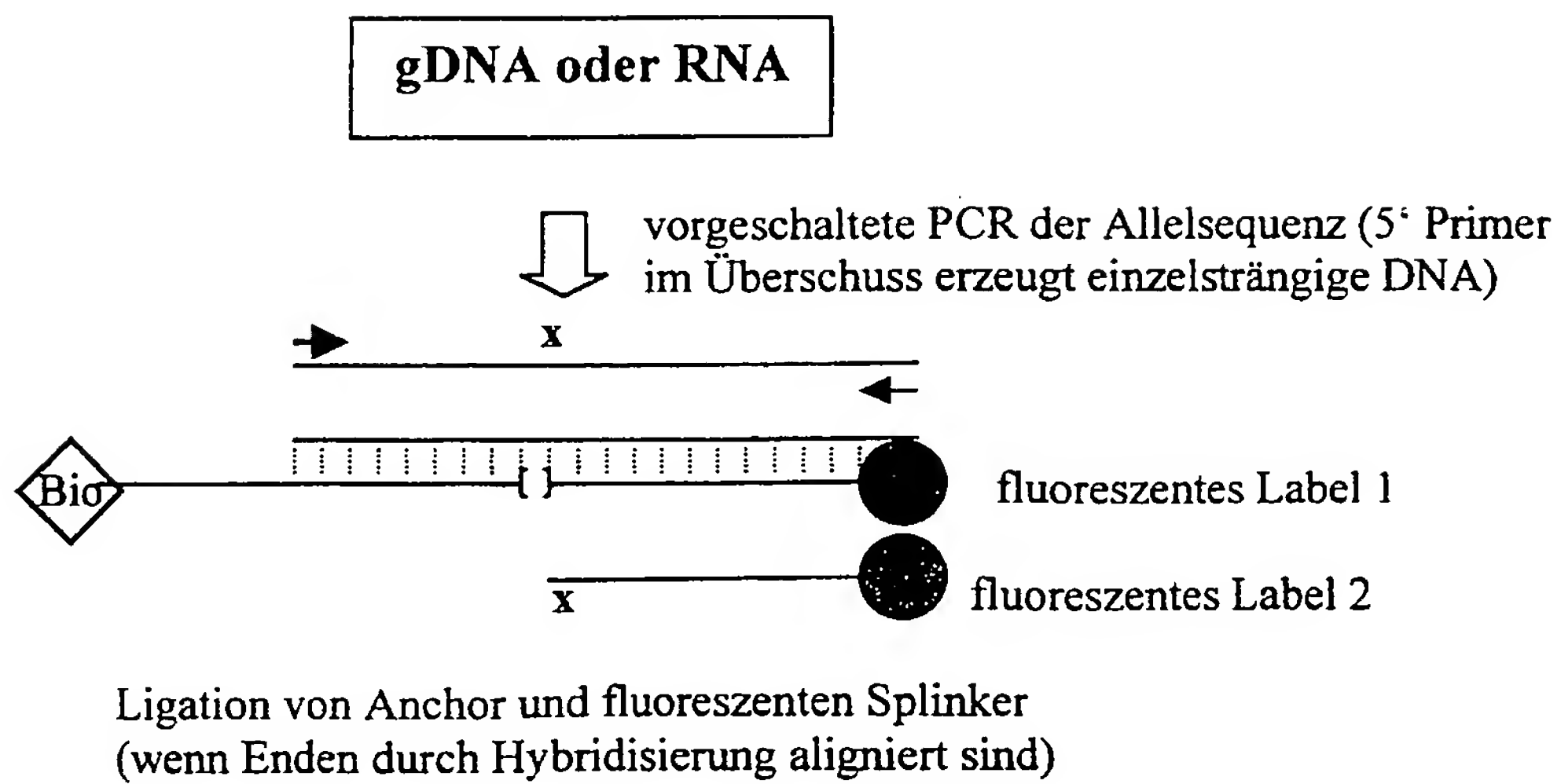
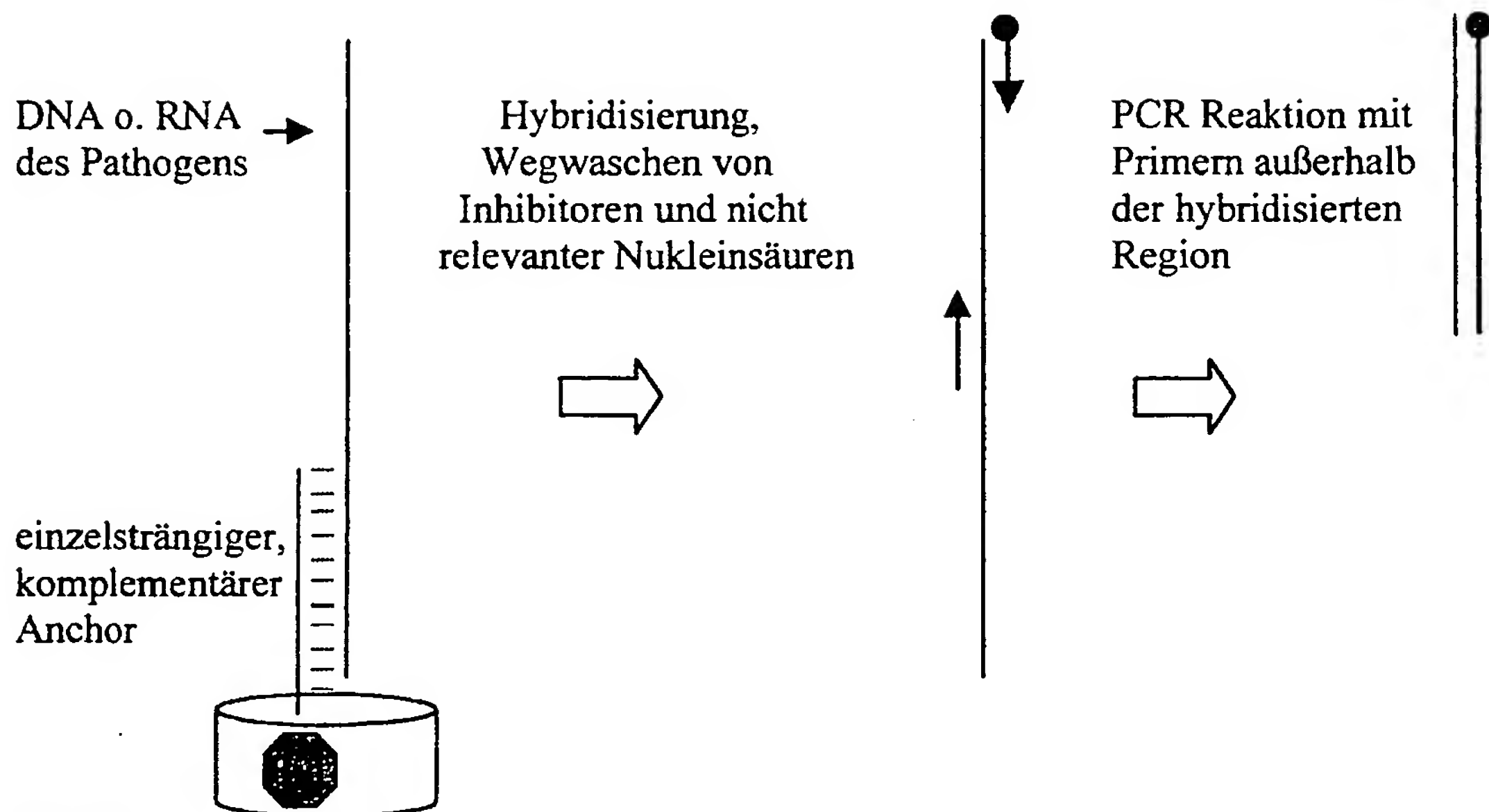


Fig. 7

5/8

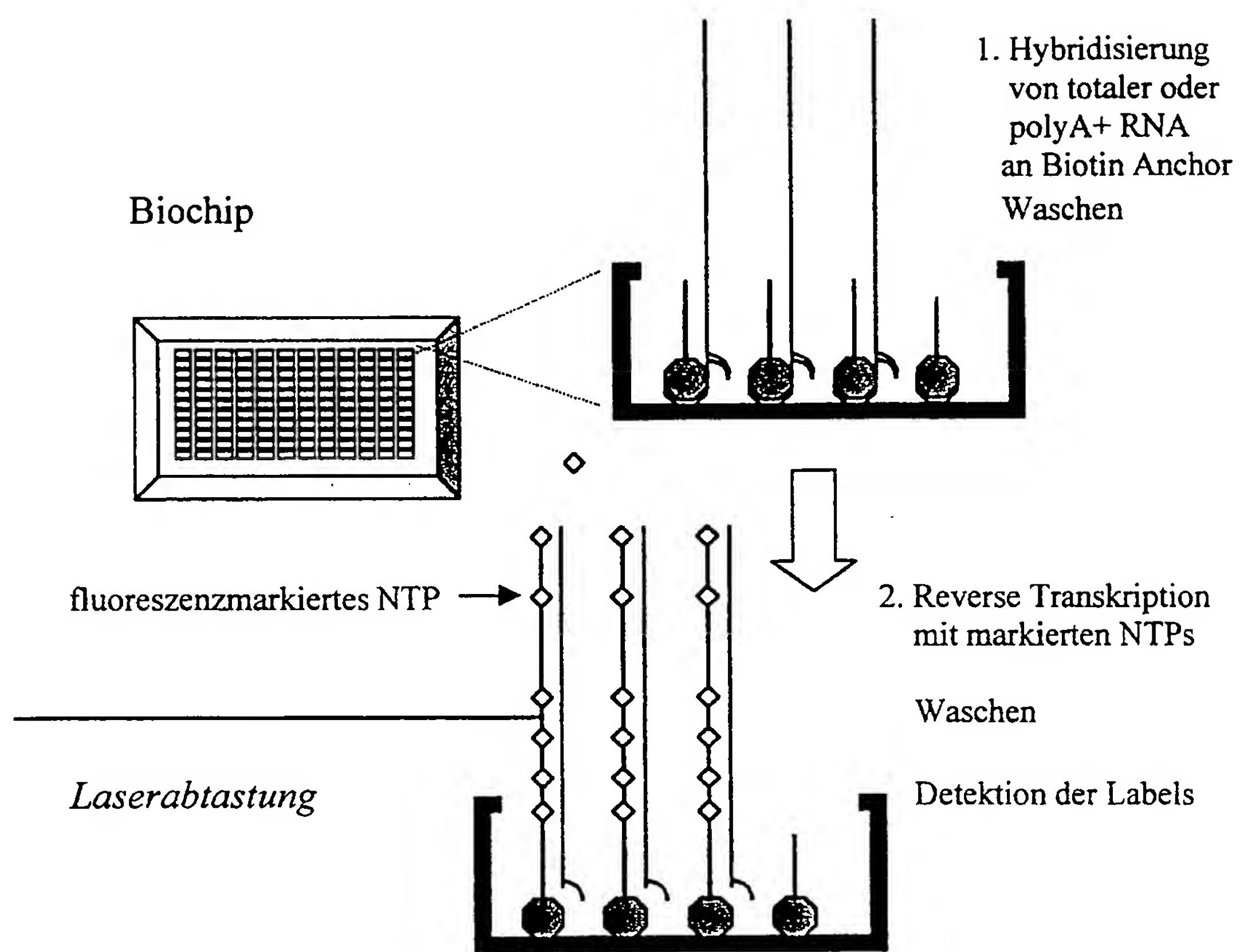


Fig. 8

6/8

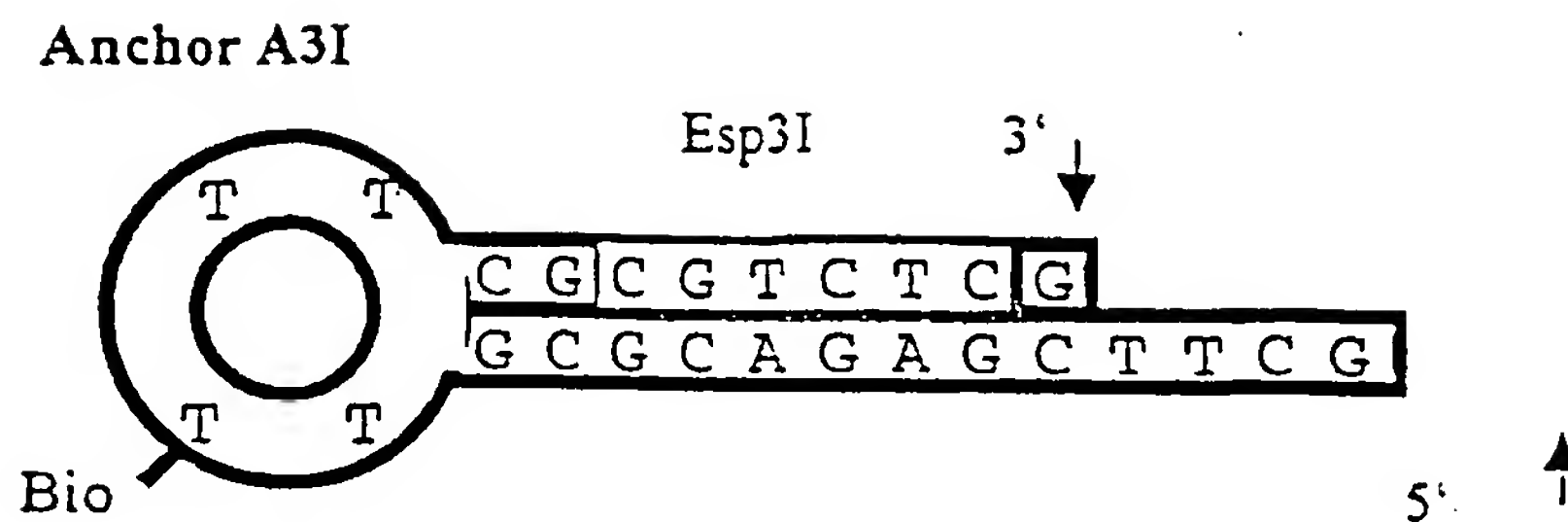


Fig. 9

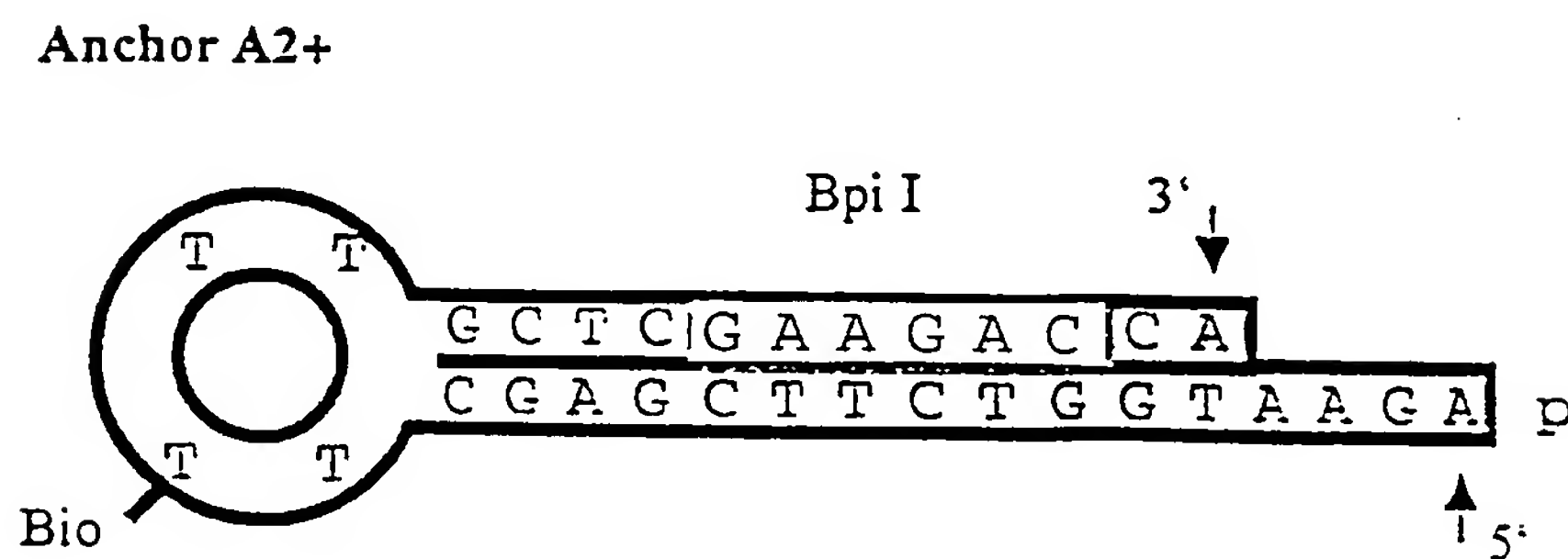


Fig. 10

Bipartite Anchor (Rekonstitution der Esp3I Schnittstelle erst nach Ligation)

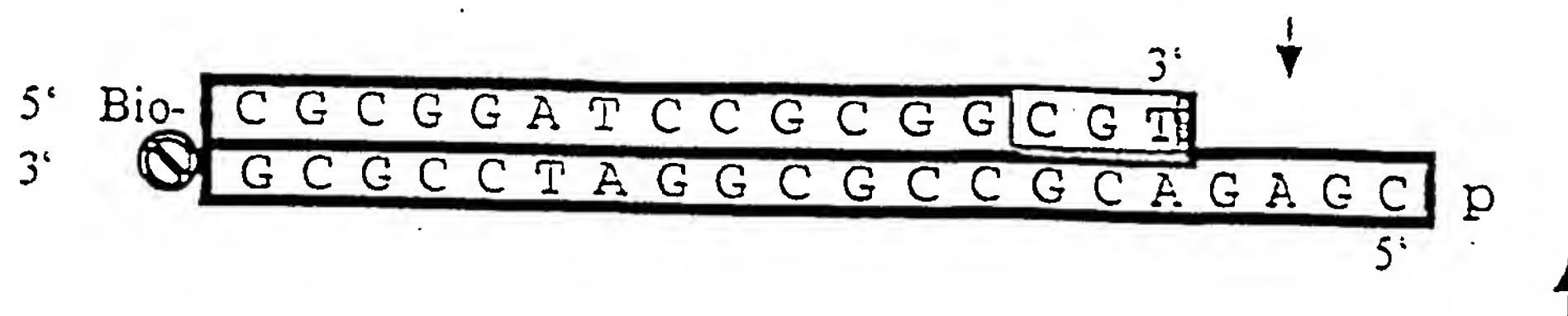


Fig. 11

7/8

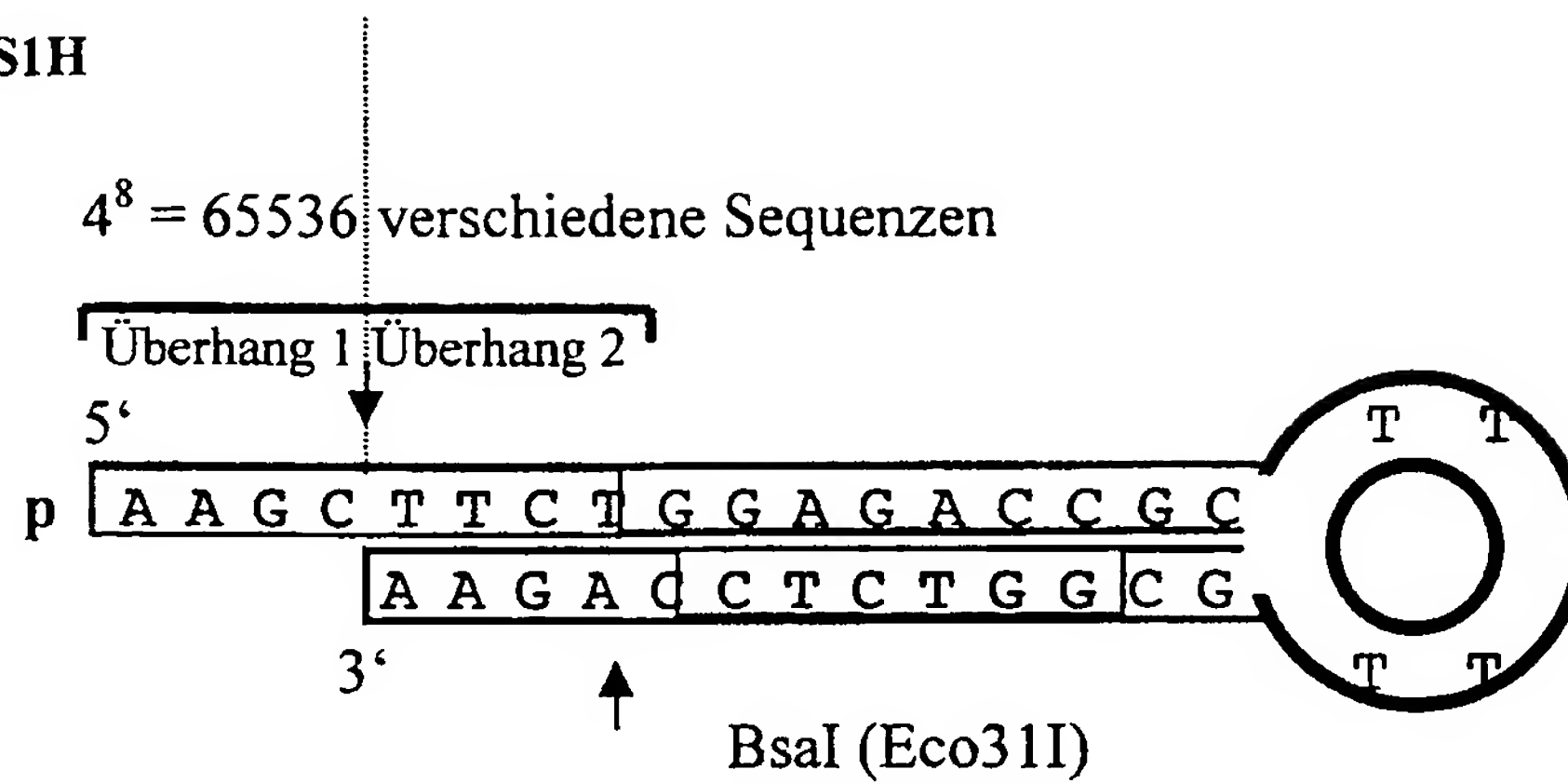
**Splinker S1H**

Fig. 12

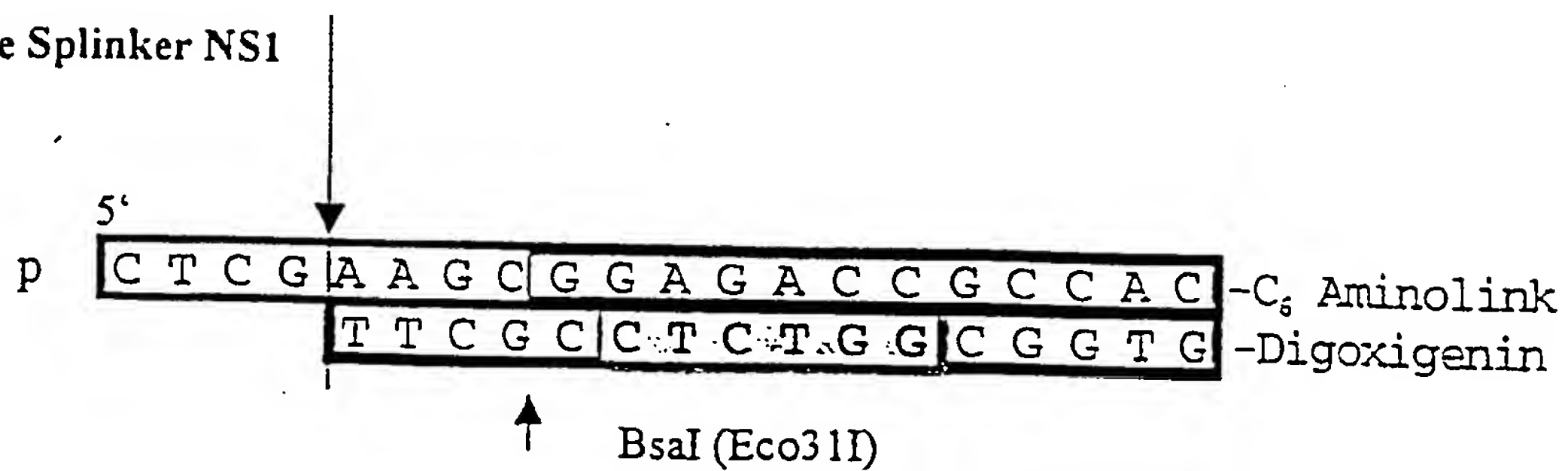
**Bipartite Splinker NS1**

Fig. 13



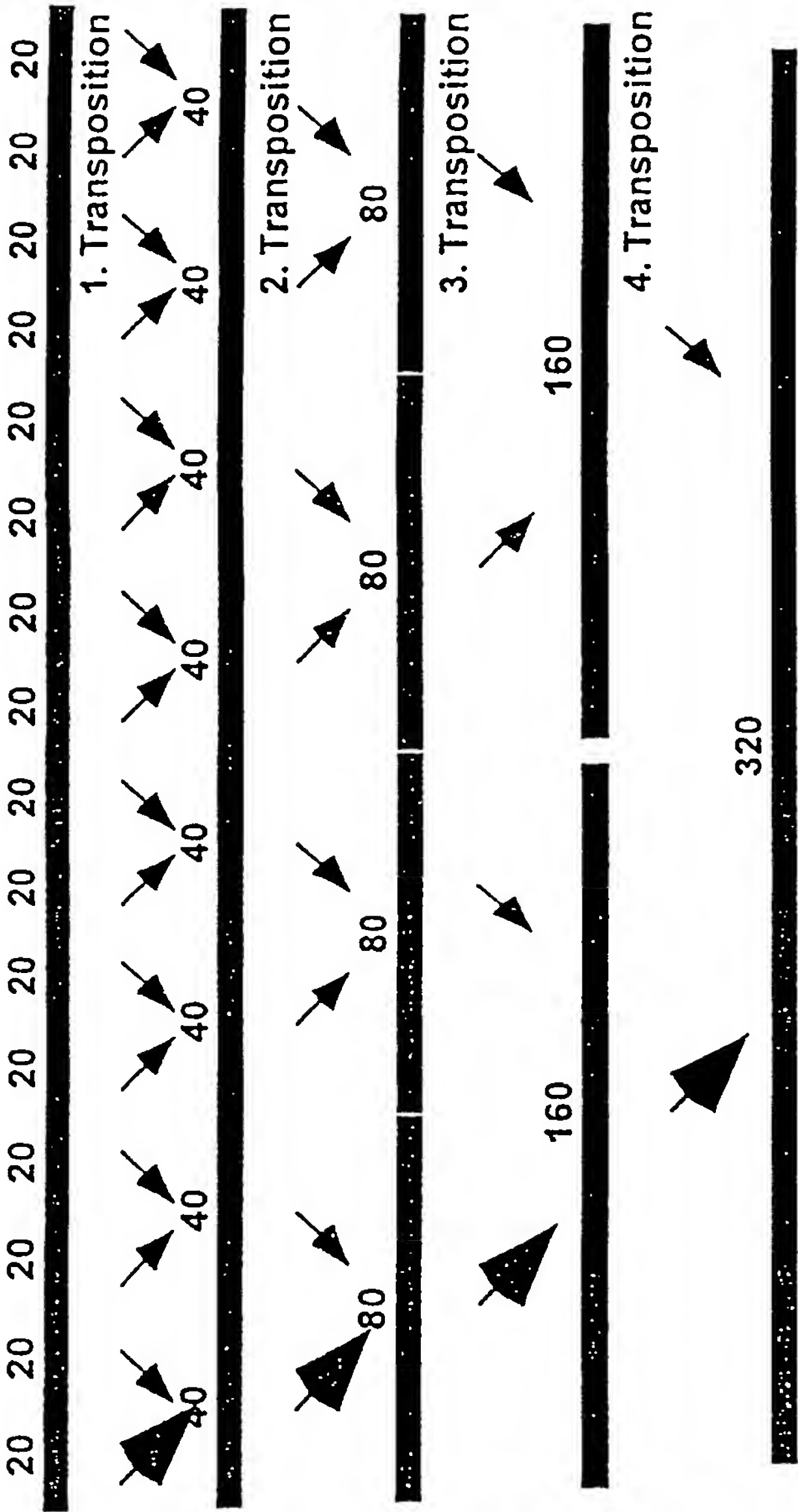


Fig. 14

## SEQUENZPROTOKOLL

<110> Diavir GmbH

5 <120> Verfahren zur Synthese von DNA-Fragmenten

<130> DV-001 PCT

10 <140> xx  
<141> 2000-06-07

<150> DE 199 25 862.7  
<151> 1999-06-07

15 <160> 7

<170> PatentIn Ver. 2.1

20 <210> 1  
<211> 26  
<212> DNA  
<213> Künstliche Sequenz

25 <220>  
<223> Beschreibung der künstlichen  
Sequenz:Oligonukleotid

30 <400> 1  
gcttcgagac gcgttttcgc gtctcg 26

<210> 2  
<211> 32  
<212> DNA  
35 <213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen  
Sequenz:Oligonukleotid

40 <400> 2  
agaatgggtct tcgagctttt gctcgaagac ca 32

45 <210> 3  
<211> 16  
<212> DNA  
<213> Künstliche Sequenz

50 <220>  
<223> Beschreibung der künstlichen  
Sequenz:Oligonukleotid

55 <400> 3  
cgcggtatccg cggcgt 16

<210> 4  
<211> 20  
60 <212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen  
65 Sequenz:Oligonukleotid

<400> 4  
cgagacgccg cggatccgcg 20

5  
<210> 5  
<211> 34  
<212> DNA  
<213> Künstliche Sequenz

10  
<220>  
<223> Beschreibung der künstlichen  
Sequenz:Oligonukleotid

15 <400> 5  
aagcttctcg agaccgcttt tgcggtctcc agaa 34

20 <210> 6  
<211> 20  
<212> DNA  
<213> Künstliche Sequenz

25 <220>  
<223> Beschreibung der künstlichen  
Sequenz:Oligonukleotid

30 <400> 6  
ctcgaagcgg agaccgccac 20

35 <210> 7  
<211> 16  
<212> DNA  
<213> Künstliche Sequenz

40 <220>  
<223> Beschreibung der künstlichen  
Sequenz:Oligonukleotid

45 <400> 7  
gtggcggtct ccgctt 16